State of The Salmon: Rapid status assessment approach for Pacific salmon under Canada's Wild Salmon Policy

Gottfried Pestal, Bronwyn L. MacDonald, Sue C.H. Grant, Carrie A. Holt

Fisheries and Oceans Canada,
Science Branch, Pacific Region
200-401 Burrard Street, 13th floor
Vancouver, BC
V6C 3S4

2023

Canadian Technical Report of Fisheries and Aquatic Sciences 3570

Canadian Technical Report of Fisheries and Aquatic Sciences

Technical reports contain scientific and technical information that contributes to existing knowledge but which is not normally appropriate for primary literature. Technical reports are directed primarily toward a worldwide audience and have an international distribution. No restriction is placed on subject matter and the series reflects the broad interests and policies of Fisheries and Oceans Canada, namely, fisheries and aquatic sciences.

Technical reports may be cited as full publications. The correct citation appears above the abstract of each report. Each report is abstracted in the data base *Aquatic Sciences and Fisheries Abstracts*.

Technical reports are produced regionally but are numbered nationally. Requests for individual reports will be filled by the issuing establishment listed on the front cover and title page.

Numbers 1-456 in this series were issued as Technical Reports of the Fisheries Research Board of Canada. Numbers 457-714 were issued as Department of the Environment, Fisheries and Marine Service, Research and Development Directorate Technical Reports. Numbers 715-924 were issued as Department of Fisheries and Environment, Fisheries and Marine Service Technical Reports. The current series name was changed with report number 925.

Rapport technique canadien des sciences halieutiques et aquatiques

Les rapports techniques contiennent des renseignements scientifiques et techniques qui constituent une contribution aux connaissances actuelles, mais qui ne sont pas normalement appropriés pour la publication dans un journal scientifique. Les rapports techniques sont destinés essentiellement à un public international et ils sont distribués à cet échelon. II n'y a aucune restriction quant au sujet; de fait, la série reflète la vaste gamme des intérêts et des politiques de Pêches et Océans Canada, c'est-à-dire les sciences halieutiques et aquatiques.

Les rapports techniques peuvent être cités comme des publications à part entière. Le titre exact figure au-dessus du résumé de chaque rapport. Les rapports techniques sont résumés dans la base de données *Résumés des sciences aquatiques et halieutiques.*

Les rapports techniques sont produits à l'échelon régional, mais numérotés à l'échelon national. Les demandes de rapports seront satisfaites par l'établissement auteur dont le nom figure sur la couverture et la page du titre.

Les numéros 1 à 456 de cette série ont été publiés à titre de Rapports techniques de l'Office des recherches sur les pêcheries du Canada. Les numéros 457 à 714 sont parus à titre de Rapports techniques de la Direction générale de la recherche et du développement, Service des pêches et de la mer, ministère de l'Environnement. Les numéros 715 à 924 ont été publiés à titre de Rapports techniques du Service des pêches et de la mer, ministère des Pêches et de l'Environnement. Le nom actuel de la série a été établi lors de la parution du numéro 925.

Canadian Technical Report of Fisheries and Aquatic Sciences 3570

2023

STATE OF THE SALMON:

RAPID STATUS ASSESSMENT APPROACH FOR PACIFIC SALMON UNDER CANADA'S WILD SALMON POLICY

Gottfried Pestal[2], Bronwyn L. MacDonald[1], Sue C.H. Grant[1], Carrie A. Holt[3]

[1]Fisheries and Oceans Canada,
Science Branch, Pacific Region
200-401 Burrard Street, 13th floor
Vancouver, BC, V6C 3S4

[2]Solv Consulting Ltd.
Unit 60607 RPO Granville Park
Vancouver, BC, V6H 4B9

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

Cat. Fs97-6/3570E-PDF ISBN 978-0-660-68325-6 ISSN 1488-5379

Correct citation for this publication:

Pestal, G., MacDonald, B.L., Grant, S.C.H., and Holt, C.A. 2023. State of the Salmon: rapid status assessment approach for Pacific salmon under Canada's Wild Salmon Policy. Can. Tech. Rep. Fish. Aquat. Sci. 3570: xiv + 200 p.

Table of Contents

Table of Tables

Table of Figures

ABSTRACT

Pestal, G., MacDonald, B.L., Grant, S.C.H., and Holt, C.A. 2023. State of the Salmon: rapid status assessment approach for Pacific salmon under Canada's Wild Salmon Policy. Can. Tech. Rep. Fish. Aquat. Sci. 3570: xiv + 200 p.

We developed an approach to rapidly assess the biological status of Pacific Salmon Conservation Units (CUs) under the Wild Salmon Policy (WSP). This approach assigns a *Red* (poor), *Amber* (intermediate), or *Green* (good) status, with *High, Medium*, or *Low* confidence, to CUs with applicable data. Existing integrated WSP status assessment approaches are labour intensive, and therefore, have only been completed for ~11% of the ~377 current Pacific Salmon CUs, and have not been updated since they were assigned in expert workshops. The WSP rapid status approach can provide a more complete and current picture of status across Pacific salmon species in BC & the Yukon. We developed seven candidate WSP rapid status algorithms based on completed WSP integrated status assessments, evaluated algorithm performance against a set of criteria, and identified the best performing algorithm. WSP rapid statuses are incorporated into DFO's new Pacific Salmon Status Scanner, an interactive data visualization tool for salmon experts. Rapid statuses in DFO's Salmon Scanner will be combined with expert review to support *Fisheries Act* Stock Management Unit (SMU) Limit Reference Point (LRP) status requirements, state of salmon reporting, climate change vulnerability assessments, and planning, prioritization and monitoring of hatchery, harvest and habitat actions.

RÉSUMÉ

Pestal, G., MacDonald, B.L., Grant, S.C.H., and Holt, C.A. 2023. State of the Salmon: rapid status assessment approach for Pacific salmon under Canada's Wild Salmon Policy. Can. Tech. Rep. Fish. Aquat. Sci. 3570: xiv + 200 p.

Nous avons élaboré une approche visant à évaluer l'état des unités de conservation (UC) de saumons du Pacifique en vertu de la *Politique concernant le saumon sauvage* (PSS). Dans le cadre de cette approche, on attribue un état *rouge* (mauvais)*, ambre* (moyen) ou *vert* (bon)*,* ainsi qu'un degré de confiance *élevé, moyen o*u *faible* aux UC selon les données applicables. Les approches intégrées d'évaluation de l'état en vertu de la PSS qui sont utilisées actuellement exigent beaucoup de travail; des évaluations de l'état ont donc seulement été réalisées pour environ 11 % des quelque 377 UC actuelles de saumons du Pacifique et elles n'ont pas été mises à jour depuis leur réalisation dans le cadre d'ateliers d'experts. L'approche rapide d'évaluation de l'état en vertu de la PSS peut fournir une représentation plus complète et à jour de l'état des différentes espèces de saumons du Pacifique présentes dans les eaux de la Colombie-Britannique et du Yukon. Nous avons développé sept choix d'algorithmes pour l'approche rapide d'évaluation de l'état en vertu de la PSS, puis nous avons déterminé l'algorithme le plus performant d'après des évaluations de l'état réalisées selon l'approche intégrée et la performance des algorithmes selon des critères établis. Les états obtenus selon l'approche rapide de la PSS ont été intégrés au nouvel explorateur de l'état des saumons du Pacifique du MPO, un outil de visualisation de données interactif destiné aux experts du saumon. Ces états intégrés dans l'explorateur du MPO seront combinés à un examen effectué par des experts afin d'appuyer les exigences en matière d'état relatives au point de référence limite d'une unité de gestion des stocks en vertu de la *Loi sur les pêches*, la production de rapports sur l'état du saumon, les évaluations de la vulnérabilité aux changements climatiques, ainsi que la planification et le suivi des mesures prises en matière d'écloserie, de pêche et d'habitat, et l'établissement des priorités connexes.

ACKNOWLEDGEMENTS

We humbly and respectfully acknowledge that this paper was written on the unceded traditional territories of the xʷməθkʷəy̓əm (Musqueam), Sḵwx̱wú7mesh (Squamish), səlilwətaɬ (Tsleil-Waututh) Nations. We recognize that these Nations have been stewards of the land and water for time immemorial, and we honour their deep understanding of the interconnectedness between people and natural systems.

We would like to thank all the DFO Staff, First Nations, NGO's, and consultants that work in the field, labs, and offices to generate salmon escapement, catch and recruitment data. The dedication, attention to detail, and hard work through all environmental conditions to collect these data is a testament to the passion and concern for salmon that so many in British Columbia and the Yukon share. These data are the foundation for all the analyses being done to support salmon resource management, including the status assessments. Algorithms can automate some steps, but without the basic data sets these WSP rapid status assessment approaches could not be conducted.

We would also like to thank all those that supported this work through the direct provision of data and many meetings and discussion on how to apply their data to the current project. This includes Joe Tadey, Matt Townsend, Chuck Parken, Antonio Vélez-Espino, Lynda Ritchie, Helen Olynyk, and Tracy Cone.

Finally we would like to thank all reviewers of this publication. This includes Chuck Parken, Antonio Vélez-Espino, Brendan Connors, Pete Nicklin, biologist for the Upper Fraser Fisheries Conservation Alliance (UFFCA), Kelsey Campbell, biologist for A-Tlegay Fisheries Society, and Ann-Marie Huang (DFO), who have supported many of the more detailed status integration processes, and have been extremely thoughtful and generous with the advice and input they provided towards this work.

FREQUENTLY ASKED QUESTIONS

Why did we develop a WSP rapid status assessment approach?

Canadian Pacific salmon abundances are broadly declining, and this is expected to continue under climate change. To track and manage these changes, it is increasingly important to have up-to-date assessments of the biological statuses for Pacific Salmon Conservation Units (CUs), which represent the fundamental units of salmon biodiversity. Annual CU status information is urgently needed to help adapt our hatchery, harvest, and habitat management approaches to changing salmon production. Existing Wild Salmon Policy (WSP) integrated status assessment processes only get us part-way there. Since they are labour and time intensive, they have only been completed for 11% of the 377 Canadian Pacific Salmon CUs and are 5-10 years out of date. The WSP rapid status assessment approach we developed enables us to assess status annually for BC and Yukon salmon CUs with applicable data.

What is a WSP rapid status assessment?

WSP rapid status assessments can assign a *Red* (poor), *Amber* (intermediate), or *Green* (good) status, with a *Low, Medium,* or *High* confidence rating, to CUs with applicable data. These statuses are generated by a computer-coded WSP rapid status algorithm, which is applied to salmon CU data. A WSP rapid status algorithm is a set of decision rules that approximate the decision-making process that experts used in WSP integrated status assessments. The algorithm assigns a WSP rapid status depending on answers to Yes/No questions for CU status data sets. The combination of metrics applied, and their individual status values compared to metric thresholds, leads to a final WSP rapid status determination.

How was the WSP rapid status assessment algorithm selected?

We developed a suite of candidate algorithms based on past WSP integrated assessments. We evaluated the performance across multiple candidate WSP rapid status algorithms, by comparing their respective rapid statuses against WSP integrated statuses assigned from past expert-driven processes. The top-performing algorithm was the *Learning Tree 3*, which we recommend for use in future applications.

What are the three core principles of the *Learning Tree 3* WSP rapid status algorithm?

The first core principle is that WSP CUs were identified and WSP rapid statuses were developed based on conservation biology principles, and scientific peer-reviewed publications. This ensures that Pacific salmon statuses are scientifically objective, consistent, and comparable across BC/Yukon CUs. Standardized metrics also need to be widely applicable and relatively easy to use and update regularly. The second core principle of WSP rapid status assessment is the vetting of data and evaluation of WSP rapid statuses by experts on specific salmon CUs. The final core principle is continual learning & refinement.

How will we use the Learning Tree 3 algorithm?

The *Learning Tree* 3 is central to DFO's Salmon Scanner. DFO's Scanner is an interactive data visualization tool specifically designed for experts to support their work on Pacific salmon. This tool enables users to compare status trends across CUs and years throughout BC and the Yukon to support scientific discovery, and decision-making processes. The Scanner will support status evaluations for stock management units against their limit-reference-points (LRP), which is a legal requirement under the modernized *Fisheries Act* (2019). Other applications could include updates on the state of the salmon, and to support climate change vulnerability assessments. In these decision-making contexts, WSP rapid statuses would be combined with expert input and peer-review.

1 INTRODUCTION

1.1 THE URGENT NEED FOR RAPID WILD SALMON POLICY (WSP) STATUS ASSESSMENTS

We present a Wild Salmon Policy (WSP) rapid status assessment approach to annually assess salmon conservation units (CUs) as *Red (poor)*, *Amber (intermediate)*, or *Green (good)* status, with *High*, *Medium* or *Low* confidence. Regular tracking of the state and distribution of salmon biodiversity is increasingly important in a changing climate. Broad declines in Canadian Pacific salmon abundances have been linked to global climate change and other factors, such as deteriorating habitats, increased fish disease, and invasive species (Grant et al. 2019).

Although adaptive diversity of Pacific salmon occurs at a range of scales that include the species, CU, population and deme, the WSP identifies diversity at the scale of CUs as fundamental units that cannot be recolonized if lost (DFO 2005; Holtby and Ciruna 2007).

DFO's WSP applies to five species of Pacific salmon: sockeye (*Oncorhynchus nerka*), Chinook (*O. tshawytscha*), coho (*O. kisutch*), pink (*O. gorbuscha*) and chum salmon (*O. keta*). DFO has the authority to manage these species and their habitat under the federal *Fisheries Act* (2019). The management of steelhead (*O. mykiss*) and cutthroat trout (*O. clarkii*) has been delegated to the Province of British Columbia and the Yukon territory, and these species are therefore not included in DFO's WSP rapid status assessments.

The WSP rapid status assessment approach builds on previously completed WSP integrated status assessments (Appendix A; Holt 2009; Holt et al. 2009; DFO 2012, 2015, 2016, 2018; Grant and Pestal 2012; Grant et al. 2020). WSP status assessments are grounded in conservation biology principles, which consider population size and trends as key components in the evaluation of conservation risk (Caughley 1994; Mace et al. 2008). The WSP status assessments also build on International Union for the Conservation of Nature (IUCN) status assessment approaches for global species (Rodrigues et al. 2006; Mace et al. 2008; IUCN 2022), which have been adopted by the Committee on the Status of Endangered Wildlife in Canada (COSEWIC's) for Canadian species (COSEWIC 2010).

The detailed WSP integrated assessments are labour intensive, taking 10-40 experts up to three days to complete for each group of CUs. For this reason, assessments have only been completed for ~11% of the 377 current Pacific Salmon CUs, and are currently 5-10 years out of date (Wade et al. 2019). The WSP rapid approach can fill gaps by expanding coverage of CU statuses across time, species, and geographic areas. The WSP rapid status approach ensures statuses are scientifically objective, consistent, and comparable across BC/Yukon CUs. This approach is also relatively easy to implement, applicable to data rich and data poor CUs, and can be updated annually.

WSP rapid statuses are included in DFO's Salmon Scanner, which is an interactive data visualization tool. DFO's Salmon Scanner is specifically designed for experts to support their work and help them contribute to scientific discovery and decision-making processes. It will be used in expert processes to track annual salmon trends, support climate change adaptive science advice to manage hatchery, harvest and habitat actions, and support *Fisheries Act* limit reference point (LRP) status evaluations for stock management units (DFO 2022).

1.2 STATUS ASSESSMENTS UNDER THE WSP

Rapid status algorithms are part of the next implementation phase of the WSP. Therefore, methods used to generate WSP rapid statuses must fit within the concepts, definitions, and practices established through 20 years of previous work (Table 1; Figures 1 and 2; Appendix A; DFO 2005, 2012, 2015, 2016, 2018; Holt 2009, 2010; Holt et al. 2009; Grant and Pestal 2012; Grant et al. 2020). The WSP rapid status approach must also remain flexible and adaptable to changes in metrics and benchmarks and lessons learned from new WSP integrated status assessments for other species and geographic areas. To set the stage for the WSP rapid status assessment work, this section presents a brief overview of key WSP concepts and terminology. Appendix A provides details that include:

- additional background on the need for standardized monitoring of Pacific salmon (WSP Strategy 1);
- details on work that has been completed to date to implement WSP Strategy 1;
- and a comparison of WSP status assessments to those completed by COSEWIC under the *Species at Risk Act* (SARA).

A CU status can fall into one of five status zones: *Red* (poor); *Red/Amber*; *Amber* (intermediate), *Amber/Green*, and *Green* (good). *Red*, *Amber*, and *Green* statuses were part of the original WSP (Figure 1; WSP 2005), while *Red/Amber* and *Amber/Green* were added through subsequent WSP integrated status assessment processes (Table 1; Grant & Pestal 2012; DFO 2015; DFO 2016; Grant et al. 2020), to represent statuses that were intermediary between *Red* and *Amber*, and *Amber* and *Green*, respectively. There is also a data deficient category (DD), for when a CU does not have sufficient data quantity or quality to assess status, and an undetermined category (UD), when an integrated status cannot be resolved.

The basic sequencing of past WSP integrated status assessment processes included:

1. Compiling CU escapement, recruitment, survival and other data for the group of CUs being assessed (Section 2.2). Spawner enumeration sites are selected, and data gaps are filled as required to assess WSP status.
2. Selecting applicable metrics for each CU's status assessments from a suite of possible metrics: *abundance, trends in abundance, fishing mortality*, and *distribution*-based metrics (Figure 2; Holt et al. 2009; Holt 2009; Holt 2010). Metric applicability depends on the type of data available for the assessed CU, and the quantity and quality of these data. Benchmarks for *abundance* metrics are estimated if possible.
3. Calculating statuses for each individual metric. Depending on the metrics used and the CU data available, not all metrics may indicate the same WSP status. For example, there were cases where a CU's *percent change* (recent three generation trend) metric might indicate a *Red* status, while a *long-term trend* metric might indicate a *Green* status (Holt et al. 2009; Grant et al. 2011; Grant & Pestal 2012; DFO 2015; DFO 2016; Grant et al. 2020).
4. Assessing CU WSP integrated statuses. In this step, experts determine WSP integrated statuses by combining individual metric statuses, with other CU information, in a workshop setting (Grant and Pestal 2012; DFO 2015, 2016; Grant et al. 2020). This generates a CU's consensus designation: WSP status; or DD; or UD. This approach was essential to developing a common rationale for considering information across different metrics in a structured and consistent way.

1.3 CORE PRINCIPLES OF THE WSP RAPID STATUS ASSESSMENT APPROACH

There are three core principles of the WSP rapid status assessment approach:

1. **The *first core principle* is that WSP CUs were identified and rapid statuses were developed based on conservation biology principles** (Mace and Lande 1991; Mace et al. 1992, 2008; Caughley 1994; National Research Council (US) Committee on Scientific Issues in the Endangered Species Act 1998; McElhany et al. 2000; Rodrigues et al. 2006), **and are aligned with scientific peer-reviewed publications (see Appendix A for more details)** (Holtby and Ciruna 2007; Holt 2009; Holt et al. 2009; Holt 2010; Grant et al. 2011; Grant and Pestal 2012; DFO 2015, 2016; Brown et al. 2019; Grant et al. 2020). This ensures that Pacific salmon statuses are scientifically objective, consistent, and comparable across BC/Yukon CUs. Standardized metrics also need to be widely applicable and relatively easy to use and update regularly. Specifically:
    - The WSP identifies diversity at the scale of CUs, as fundamental units that cannot be recolonized if lost (DFO 2005). Methods to identify CUs (Holtby and Ciruna 2007), revisions (Grant et al. 2011; Brown et al. 2019), and process to revise CUs (Wade et al. 2019) have been peer reviewed through DFO CSAS (see Appendix A.2).
    - The CU WSP integrated status assessments that have been completed were based on ~15 years of development, including CSAS meetings that took up to 3 days and 40 experts to complete (Holt 2009; Holt et al. 2009; Grant et al. 2011; Grant and Pestal 2012; DFO 2015, 2016; Grant et al. 2020). These processes were collaborations across CU experts including DFO staff, Indigenous community members, consultants, NGO's, etc. (see Appendices A.3 & A.4).
    - The WSP status assessment approach builds on the status assessment approach used by the International Union for the Conservation of Nature (IUCN) for global species (Rodrigues et al. 2006; Mace et al. 2008; IUCN 2017), which has been adopted by the Committee on the Status of Endangered Wildlife in Canada (COSEWIC's) for Canadian species (COSEWIC 2010). A COSEWIC species is roughly equivalent to a WSP CU. The WSP *Red* status zone largely aligns with COSEWIC's *Endangered* status; *Amber* aligns with *Threatened* and *Special Concern* status; and *Green* aligns with *Not at Risk* status (See Appendix A.5).
2. **The *second core principle* of WSP rapid status assessment is the vetting of data and evaluation of WSP rapid statuses by CU experts**. DFO stock assessment leads work in collaboration with Indigenous groups, consultants, and others that support or lead salmon stock assessment programs. These CU experts work iteratively to fine tune the CU data used (determining appropriate escapement sites, years, data treatment, etc.), and select applicable WSP status metrics and metric calculation details, given their knowledge of the data.
3. **The *final core principle* of the WSP rapid status algorithm is continual learning and refinement.** This means that data sets and status metrics for each CU will be regularly reviewed and updated, and that the rapid status algorithm will be reviewed through on-going work with CU experts (described in the second core principle). By evaluating WSP rapid status algorithm outputs for the CUs for which they have

expertise, CU experts can identify where decision rules may be revised or added to the WSP rapid status algorithm. As new CU cases are added, where common new and/or revised decision-rules are proposed, the revised WSP rapid status algorithm performance can be tested for overall improvements.

1.4 KEY TERMINOLOGY FOR THE WSP RAPID STATUS ASSESSMENT APPROACH

Throughout this report several terms are used with specific definitions.

- **Metrics:** quantitative metrics developed for WSP status assessment: *relative abundance; absolute abundance; long-term trend in abundance; and percent change (short-term trend in abundance)*(Section 2.2.2).
- **Benchmarks**: specific values identified under the WSP to delineate between *Red, Amber*, and *Green* status zones for each metric. For example, 50% is the lower benchmark for the *long-term trend* metric, that delineates this metric's *Red* and *Amber* status zones (Section 2.2.2). This metric compares the ratio of the current generational average (geometric mean) spawner abundance to the long-term average (geometric mean) to lower (50%) and upper (75%) benchmarks.
- **WSP rapid status algorithms:** sets of decision rules that approximate the decision making process that experts used in WSP integrated status assessments (Section 2.4); see Figure 20 as a candidate algorithm used in the WSP rapid status process.
- **Performance measures:** summary statistics used to compare the performance of candidate algorithms (e.g. number of correct WSP rapid statuses across CU cases, compared to WSP integrated 'true' statuses for identical data) (Section 2.3).

1.5 REPORT OUTLINE

The objective of this paper is to present a WSP rapid status approach for Pacific Salmon CUs. This is central to DFO's new Pacific Salmon Status Scanner, which is an interactive data visualization tool designed for experts. Experts include research scientists, Indigenous technical experts, stock assessment biologists, habitat, harvest, and hatchery management biologists etc. This paper provides the following:

- Methods & Results that include:
    - the *learning* and *retrospective (out-of-samples) data sets* used;
    - development of candidate WSP rapid status algorithms;
    - the performance evaluation of rapid status algorithms, which uses various performance metrics to compare rapid statuses to 'true' WSP integrated statuses from previous assessments; criteria are identified to evaluate results;
    - two sensitivity tests: an evaluation of past annual WSP rapid statuses produced for a CU: *retrospective (out-of-samples)* test; and a comparison of CUs rapid statuses with and without their *relative-abundance* metric;
    - capturing confidence in WSP rapid status designations;
- Discussion: selected algorithm, future considerations, and potential applications.

2 METHODS

2.1 ANALYSIS OUTLINE

To develop a rapid status algorithm that approximates the detailed WSP integrated status assessment approach, we worked through the following 11 steps:

1. Compiled the *learning data set.* This includes metric values, corresponding metric statuses, and WSP integrated statuses (considered ‘true’ statuses) for CUs from the four past WSP status assessment processes (Grant and Pestal 2012; DFO 2015, 2016; Grant et al. 2020) (Section 2.2).
2. Identified six performance criteria to guide the construction, evaluation, and selection of candidate algorithms. Performance evaluation included quantitative error evaluation, and qualitative considerations (Sections 2.3 & 2.4)
3. Fit Classification and Regression Tree (CART) models to the learning data set: this includes metric values or statuses and corresponding WSP integrated statuses (‘true’ statuses) derived from existing WSP integrated status assessments. Trees were fit using various combinations of predictor variables (metric values and statuses on those metrics) and response variable formats (CU status), data subsets, and model fit settings (e.g. complexity penalty; error weighting).
4. Selected candidate CART algorithms (‘fitted algorithms’) to span the range of trees possible from the available data and settings, from very simple to very complex.
5. Reviewed narratives provided by experts for their CU WSP integrated status designations, in order to extract common rationale for these designations across CUs: these narratives are reprinted in Appendix B from published (Grant and Pestal 2012; Grant et al. 2020), and unpublished reports (Brown et al. 2014; Parken et al. 2014).
6. Developed custom algorithms (‘constructed algorithms’) by combining CART-derived algorithm branches from step 4, with common rationale from WSP integrated status assessments in step 5.
7. Implemented candidate algorithms as an R function to estimate WSP rapid statuses using existing WSP status assessment metrics (*learning data set*).
8. Evaluated algorithm performance according to the criteria identified in Step 2.
9. Conducted *retrospective* (*out-of-samples*) tests using the selected algorithm on years that do not have WSP status assessment completed.
10. Reviewed results with salmon stock assessment experts.
11. Performed sensitivity tests (with and without using *relative abundance* metrics).

Steps 3 and 4 above were repeated through an iterative, collaborative process as we explored the effect of alternative CART settings and identified a shortlist of candidate CART algorithms.

Steps 5-10 above were also repeated through an iterative process. Constructed algorithms were developed and refined through evaluating performance and reviewing documentation from the status workshops to identify missing components and uncover special considerations.

2.2 DATA

2.2.1 Two Data Sets: Learning vs. Retrospective (Out-of-Samples)

We used two data sets for the development and performance evaluations of WSP rapid status algorithms: *the learning data set* and the *retrospective (out-of-samples) data set.* The key differences between these data sets are that the *learning data set* used the exact data and metric values from past WSP integrated status assessments, and the *retrospective (out-of-samples) data set* used the latest available data and metrics at the time of this publication, for each CU (details below).

2.2.1.1 Learning Data Set

The first phase of algorithm development was to build a *learning data set.* The purpose of the *learning data set* was to support the development and evaluation of the candidate algorithms.

The *learning data set* for this analysis consists of WSP metric values, metric statuses, and corresponding WSP integrated statuses from the four completed WSP integrated status assessments (Appendices B & C). This included two assessments for Fraser sockeye (Grant et al. 2011; Grant and Pestal 2012; Grant et al. 2020), one for Interior Fraser coho (Parken et al. 2014; DFO 2015), and one for Southern BC Chinook (Brown et al. 2014; DFO 2016; Brown et al. 2019).

The *learning data set* includes 65 cases for which integrated statuses were assigned by experts in the WSP integrated status assessment workshops: 22 Fraser sockeye CUs from the first status assessment, 23 Fraser sockeye CUs from the reassessment, five Interior Fraser coho CUs, and 15 Southern BC Chinook CUs.

- There are more cases than CUs, since Fraser sockeye (24 CUs in total) had two WSP integrated status assessments completed and three cases were excluded, totalling 45 cases for this CU group (Table 5). Chilko-ES was excluded for both assessments since it was rolled up into a merged Chilko-S/Chilko-ES CU due to data issues. Spawner estimates for the two CUs could not be separated at the time of assessment. Since Chilko-S contributes the most to the abundance of this pair of CUs, the WSP status will be more representative of this CU, while the Chilko-ES CU is considered data deficient (Grant et al. 2011). The first WSP integrated status assessment of Seton-Late was excluded because its status was assigned undetermined (UD) by experts (Grant & Pestal 2012).
- There are five cases for five Interior Fraser coho CUs (DFO 2015) (Table 5).
- There are 15 cases for 15 Southern BC Chinook CUs (Table 5). WSP integrated status assessments were completed for the wild component of 15 Southern BC Chinook CUs, using data from persistent survey sites classified as low or unknown enhancement (DFO 2016).

2.2.1.2 Retrospective (Out-of-Samples) Data Set

A *retrospective data set* was also built to support the evaluation of candidate algorithms for all applicable years in each CU's time series. The purpose of the *retrospective data set* was to produce an *out-of-samples* data set, where previous WSP integrated status assessments have not been completed. This was done to examine the applicability of the algorithm to new years. The *retrospective data set* also serves as an update to the data used in the status workshops, so that we can evaluate whether status has likely changed for a CU since the WSP integrated assessments were completed 5-10 years ago.

This data set used the most up to date data available at the time of this publication. This can include revisions to historical numbers, and changes in approach. Therefore, the metrics calculated for the *retrospective data set* do not align exactly with those in the *learning data set* for the same year.

The most up-to-date escapement data sets available were obtained for each CU from DFO Stock Assessment experts (Fraser sockeye: T. Cone; Interior Fraser coho: L. Ritchie; Southern BC Chinook: L.A. Vélez-Espino). CU data were prepared using the same methods used in the WSP integrated status processes including spawner abundance site selection, infilling, and, considerations of hatchery abundance proportions (Grant et al. 2011; Grant & Pestal 2012; Grant et al. 2020; Brown et al. 2014; Parken et al. 2014). We then applied the *WSPMetrics* R package (Pestal and Holt 2020) to calculate the key metrics for the time period 1995-2019, or starting three generations after the first data point for shorter data sets.

The *retrospective data set* covers a total of 42 CUs: 22 Fraser sockeye CUs (combines Chilko-S/ES and excludes Cultus-L), five Interior Fraser coho CUs, and 15 Southern BC Chinook CUs. Differences in the number of CUs covered are due to reviews or reconsiderations of the available data and, for Chinook, updated classifications of hatchery abundance proportions for component sites. Specifically, for each of these groups we provide the differences between the *retrospective time series* compared to the *learning data set*, where these occur:

<u>Fraser Sockeye:</u> There were no changes in the CU list or data treatment for 23 out of the 24 CUs between the *retrospective* and the *learning data set*. Note that similar to the *learning data set,* Chilko-ES/S were combined in the *retrospective data set,* dropping the total CUs from 24 to 23. Data for 2016-2019 were added to the *retrospective data set* using consistent data treatment methods (Grant et al. 2020). Cultus-L was excluded from the *retrospective data set* due to high hatchery contributions, and, therefore, it is not considered 'wild' according to the WSP (DFO 2005).

Salmon are considered 'wild' if they have spent their entire life cycle in the wild and originate from parents that were also produced by natural spawners that continuously lived in the wild' (DFO 2005). In the WSP integrated status assessments (Grant et al. 2011; Grant & Pestal 2012; Grant et al. 2020), adipose-fin clipped adults were removed from the escapement time series, since they represent hatchery origin fish. A significant proportion of returning adults, however, came from parents that themselves were hatchery origin, therefore, these fish would not be considered 'wild' according to the WSP. Since considerations of how to consider hatchery origin fish in this time series are outstanding at the time of this publication, we excluded this CU in the *retrospective data set*. This contrasts with the *learning data set* where Cultus-L was included because it was assessed in past WSP integrated status assessments.

<u>Interior Fraser Coho</u>: There were no changes in the CU list and data treatment between the *retrospective* and *learning data set*. Specifically, there were 5 cases for Interior Fraser coho (DFO 2015) included both data sets.

<u>Southern BC Chinook</u>: There were no changes in the CU list, and some data treatment changes between the *retrospective* and *learning data set,* based on an unpublished data report (Brown et al. 2014) a more recent published report (Brown et al. 2019), and updates provided by DFO Area staff in the summer of 2023. There were 15 cases for Southern BC Chinook in the *retrospective data set*, with some differences from the *learning data set.* These

differences are due to the more recent reclassification of some Southern BC Chinook sites, and a review of site-specific classifications of enhancement level (Brown et al. 2020). As a result, four of the original 15 CUs can no longer be assessed: a) because the data were determined to be too poor for calculating metrics; and b) because there are now no wild sites. However, four additional CUs can be assessed, two because survey sites were re-classified from high or moderate to low enhancement (Lower Shuswap River in the Shuswap River_SU_0.3 CU, CK-15, Marble and Cayeghle in the West Vancouver Island-North_FA_0.x CU, CK-33) , one because data quality was reassessed (Lower Fraser River_SP_1.3, CK-04), and one because new data was provided by area staff for a CU that was previously data deficient (Okanagan_1.x, CK-01). The *retrospective data set,* therefore, covers 15 CUs of Southern BC Chinook.

2.2.1.3 Annual Updates

The data processing procedures and R code we developed to create the *retrospective data set* also set the stage for annual updates and expansion to additional CUs. This includes code and other inputs developed to clean, infill and merge CU data, and calculate the annual metrics. This process also generates required input files for DFO's Salmon Scanner, which allows users to interactively analyze statuses over time, as well as across species and areas (Section 4.5). These data processing procedures and R code will be used to update these time series annually, and may evolve as required by data analysts.

2.2.2 Overview of the WSP Rapid Status Metrics and Benchmarks

The WSP emphasizes 'standardized monitoring of [Pacific] salmon status' (DFO 2005; Holt et al. 2009). A standard suite of metrics is foundational to assessing 'wildlife species' status (Rodrigues et al. 2006). Standardized status metrics have been established globally through the IUCN (Mace et al. 2008; IUCN 2022) and adopted in Canada by COSEWIC (COSEWIC 2010). Wildlife species assessed by the IUCN include plants, animals and fungi; and species assessed by COSEWIC include native mammals, birds, reptiles, amphibians, fish, arthropods, mollusks, vascular plants, mosses and lichens. Standardized metrics enable the objective and transparent assessment of status, and the production of consistent status results. A COSEWIC 'species' largely aligns with WSP CUs. Standardized metrics enable comparisons of status across assessed species and CUs. Standardized metrics also need to be widely applicable and relatively easy to use and update regularly (Mace and Lande 1991).

The WSP status assessment approach was built on the status assessment methods developed by the IUCN and COSEWIC (DFO 2005; Holt et al. 2009). The WSP status approach currently includes metrics for a CU's abundance and trends in abundance, described in subsequent sections (Appendix A; Holt et al. 2009; Holt 2009; Grant et al. 2011; Grant & Pestal 2012; DFO 2015; DFO 2016; Grant et al. 2020). Status criteria are based on conservation biology principles, which emphasizes two paradigms: small population size and declining population to assess conservation risk (Caughley 1994; Mace et al. 2008).

Other related information on salmon CUs such as spawner fecundity, size-at-age, total productivity or survival at different life-stages has been considered in WSP integrated status assessments, directly through the data sets used, or in final status narratives.

While there is a considerable amount of other ancillary information such as information on salmon disease, parasite prevalence, genetic diversity, etc. that could be included to assess salmon status, we do not recommend using these sources of information for the WSP rapid status approach. Instead we recommend continuing to emphasize standardized metrics and additional information that focuses on abundance and trends in abundance. It would also be challenging to implement these in a standardized way across CUs to assess WSP status due to their limited availability across CUs, and gaps in determining thresholds to distinguish between poor, intermediate and good statuses for this information. Where this type of ancillary information would be particularly important, however, is to help understand threats faced by a salmon CU, and, therefore, support rebuilding considerations.

The following four metrics were applied in WSP status assessment processes conducted to date: *relative abundance, absolute abundance, long-term trend* and *percent change*. The *learning data* set and the *retrospective (out-of-samples) data* set include all four of these metrics, where available for a CU. We also discuss *distribution* metrics and ancillary information used specifically in past WSP integrated status assessments, and rationale for exclusion in the WSP rapid status approach. Specifics on these metrics are provided below:

Relative abundance

This metric compares a CU's most recent generational average spawner abundance to benchmarks estimated with a) stock-recruitment models (Holt et al. 2009; Grant et al. 2011; Grant and Pestal 2012), b) freshwater habitat capacity models (Parken et al. 2006; Grant et al. 2011; Grant and Pestal 2012; DFO 2015, 2016; Grant et al. 2020), or c) percentiles of the spawner abundance time series (Holt et al. 2018). These benchmarks are unique to each CU. Across all estimation approaches, the *relative abundance* metric is applied only when CU experts both confirm its applicability to the existing CU data, and provide a benchmark they consider appropriate.

Stock-recruitment-based benchmarks are recommended for CU's with applicable stock-recruitment data. Using this method, the lower benchmark is S_{gen}, the escapement that would result in recovery to S_{msy} in one generation, and the upper benchmark is 80% S_{msy}, which is the spawner abundance at maximum sustainable yield (Holt 2009, 2010; Holt et al. 2009). When used for southern BC Chinook CUs, the upper benchmark differed slightly (85% S_{msy}) to align with those used in the Pacific Salmon Treaty (PST) process.

For CUs where stock-recruitment data are not available, habitat-capacity-derived benchmarks have been used for the *relative-abundance* metric. For lake-rearing sockeye CUs, habitat capacity based on the rearing lakes used by a CU's juvenile stages in freshwater have been used. The lake(s) photosynthetic rate (PR) and juvenile sockeye competitors (Grant et al. 2011) are used to estimate S_{max}, which are spawners at maximum juvenile production. The recommended lower and upper benchmarks are respectively, 20% and 40% of S_{max} (Holt 2009; Grant et al. 2011). For a number of Chinook CUs, freshwater habitat capacity has been used to develop relative-abundance-metric lower (S_{gen}) and upper benchmarks (85% S_{msy}) (Parken et al. 2006).

Percentile benchmarks have been recommended for data limited CUs. However, they have been shown to be appropriate only in cases where CU productivity is moderate to high (>2.5 recruits-per-spawner) and harvest rates are moderate to low (≤ 40%) (Holt et al. 2018). They have not yet been used in WSP integrated status assessment processes, and since these were not provided by any CU experts for this current WSP rapid status assessment process, they are not included here.

Absolute abundance

This metric compares the average escapement of the most recent generation (geometric mean) to COSEWIC criterion D1 and part of criterion C, which are used to define 'Threatened Species' (COSEWIC 2020). The lower benchmark is set at 1,000, to align with criterion D1, and the upper benchmark is set at 10,000, which is used in combination with other abundance metrics under criterion C (COSEWIC 2020).

These benchmarks are grounded in fundamental principles of population and conservation ecology. The value 1,000 is a critical threshold identified in conservation biology (National Research Council (US) Committee on Scientific Issues in the Endangered Species Act 1998; McElhany et al. 2000). Below 1,000, a population is more at risk from demographic stochasticity, such as randomly in a given year producing mostly males or females. They also are at greater risk from environmental change and catastrophic events, accumulating deleterious genetic mutations, and have a low evolutionary potential to adapt to environmental change.

The value 10,000 is an upper limit on population size conservation risk from environmental variation and catastrophic events; sizes above 10,000 individuals protect populations from moderate to high environmental variation as one example (National Research Council (US) Committee on Scientific Issues in the Endangered Species Act 1998; McElhany et al. 2000). Currently, deteriorating environmental conditions are increasingly occurring in salmon habitats due to climate change, with more extreme events like flooding, drought, fires, and heatwaves (Bush and Lemmen 2019; Cheung et al. 2021; IPCC 2022a; Cheng et al. 2023). These events can also occur concurrently, compounding their impacts on wildlife species.

For these reasons, the IUCN, and COSEWIC also include a small population size criteria to account for this increased extinction risk within their status assessment process (COSEWIC 2010; IUCN 2022). Wildlife species assessed by these organizations may be perpetually classified in *Threatened* or *Endangered* categories. Conservation science shows that higher extinction risk exists for such small populations regardless of whether they have remained stable at low abundances for several generations.

This 1,000 benchmark was used by experts in the past WSP integrated status assessments, in combination with other metrics and additional information, to determine CU status documented in the narratives for Fraser sockeye, Southern BC Chinook and Interior Fraser coho (see Appendix B for narratives reprinted from CSAS publications).

Long-term trend in abundance

This metric compares the ratio of the current generational average (geometric mean) spawner abundance to the long-term average (geometric mean) to lower (50%) and upper (75%) benchmarks.

Percent change (short-term trend in abundance)

This metric compares the linear change in total spawner abundance (or effective female spawners for Fraser sockeye CUs) over the most recent three generations to lower (-25%) and upper (-15%) benchmarks.

Distribution Metrics

An additional class of metrics summarizing *spawner distribution* was included in the WSP status assessment toolkit (Holt et al. 2009), but was not included in the WSP rapid status assessment approach. Further, no benchmarks have been resolved for *distribution* metrics through expert processes or research.

Distribution metrics were only applied in WSP integrated status assessments for Southern BC Chinook and Interior Fraser coho, and they did not consistently influence statuses where they were considered (DFO 2015, 2016). For other species like Fraser sockeye, they were not considered applicable. This was due to the relatively small spatial distribution of spawners within a CU relative to other species (Grant et al. 2011), with sockeye CUs generally being defined based on rearing lake and timing. Further, common sockeye assessment methods such as using fences, mark recaptures, or sonar, do not provide readily available data on spawning distribution to assess this. These stock assessment methods provide single escapement estimates for an entire system, rather than spawning locations within the CU.

Determining if salmon abundance data are a suitable proxy for changing spawner distributions should be investigated for WSP status assessments. *Distribution* metrics might be particularly important to broadly distributed CUs, like those of chum and pink salmon. If work is done to develop benchmarks and explore their use by experts in WSP integrated status assessment processes, this metric could be added to the WSP rapid status approach. However, another important consideration is how broadly available these data will be across CUs, and how readily they can be updated annually.

Information on changes within a CU's spawning or juvenile rearing distribution should be captured when developing recovery or rebuilding plans. Considerable information on spawning distribution exists among salmon experts within DFO and among Indigenous communities and other groups. This might be more relevant at this subsequent step, rather than in the evaluation of status, where abundance and abundance trends could potentially be used as a proxy for changes in distribution over time.

Ancillary Information used in past WSP integrated status assessment processes

Additional information that supported WSP integrated status assessments but was excluded from the WSP rapid assessments included: time series plots of CU escapements, recruitment, productivity, marine and/or freshwater survival, individual population escapements within a CU, exploitation rates, and hatchery information wherever relevant and available. It also included additional context for the metrics themselves, such as retrospective metric values, and uncertainty in the metrics, or, in the case of the *relative abundance* metric, uncertainty in the benchmark estimates (S_{gen} and S_{msy}). This information was included alongside the WSP metrics in CU data summaries, which were used in each of the WSP integrated status assessment processes. However, since the interpretation and use of this information varied by CU, as well as among expert groups in the workshop setting, it could not be used to inform the WSP rapid status approach.

2.2.3 WSP Rapid Status Metric Calculations

All four metrics (*relative abundance, absolute abundance, long-term trend,* and *percent change*) incorporate an estimate of the generational average of spawner abundance in their calculation. These calculations differed among groups of CUs. The generational average is calculated as the geometric mean across the number of years corresponding to the most common age class (e.g., four years for most Fraser sockeye). For Fraser sockeye CUs, spawner time series were smoothed prior to calculating generational averages, whereas for Interior Fraser coho and Southern BC Chinook generational averages were calculated using unsmoothed time series, in part because of high proportions of missing data that made generational smoothing unreliable (SBC Chinook).

2.2.4 WSP Rapid Status Metrics and Metric Values Applied

Each of the four metrics cannot be applied to every CU. The two most broadly applicable metrics are the *long-term trend* and *percent change*, since they can be applied to both types of escapement data available for Pacific salmon that include: absolute abundance or indices of abundance (i.e. relative index). Use of these metrics, however, requires a sufficient recent time series, with limited gaps. In contrast, *relative* and *absolute abundance* metrics are less broadly applicable across CUs. They require absolute abundance data, which is available for a smaller proportion of CUs.

Metric availability by CU was handled differently in the *learning data set*, which focused on the information used for past WSP integrated status assessments, and the *retrospective data set*, which captures our current understanding of best available information.

2.2.4.1 Learning Data Set Metrics

The *learning data set* included the exact metric values that were used in the past WSP integrated status assessments (Pestal & Grant 2012; DFO 2015; DFO 2016; Grant et al. 2020). This means that we did not re-calculate metric values with updated data sets and/or recalculate CU-specific benchmarks for the *relative abundance* metric. Published WSP metric values and statuses, and the resulting WSP integrated statuses from the past four integrated WSP integrated status assessment processes are listed in Appendix C, filtered according to the conditions listed below. We also removed cases from the *learning data set* where the WSP integrated status is DD or UD; this results in a total of 65 cases of combined metrics and WSP statuses. Note that the majority of cases in the *learning data set* are Fraser sockeye CUs (45/65 = 69%), because those WSP status assessments covered a lot of CUs, and it is the only CU group where a second status assessment was completed (Grant et al. 2020).

Relative abundance metric benchmarks included in the *learning data set* include stock-recruitment model benchmarks, used for most Fraser sockeye and Interior Fraser coho CUs, and one Southern BC Chinook CU (Lower Fraser River_FA_0.3); see benchmark nuances for cyclic and non-cyclic Fraser sockeye CU in subsequent paragraphs. This also included lake-carrying-capacity benchmarks for one Fraser sockeye CU (Chilliwack-ES).

Although habitat-capacity approaches were used to estimate *relative abundance* metric benchmarks for a number of Southern BC Chinook CUs in past WSP integrated status assessments (Parken et al. 2006), these were not used by experts to determine CU status. Therefore, these benchmarks were excluded from the *learning data set* because the participating experts considered them not applicable as calculated at the time. This was because most of the Chinook spawner time series used to compare recent abundances to the benchmarks represented index systems only (not the total CU), while benchmarks were calculated based on the freshwater capacity of entire CUs.

Metric values used in the completed WSP integrated status assessments had to be transformed for the algorithm inputs. This is because algorithm fitting requires that a single value and status be identified for each of the four metrics used to assess WSP rapid statuses:

- *Trend metrics*: The *long-term trend* metric values were originally expressed as a ratio (e.g.1.28 if the current generation average is 128% of the long-term average) and the *percent change* metric (previously called the short-term trend metric) values were originally expressed as a percent (e.g. -40 for a 40% decline). For consistency in the WSP rapid status assessment approach, both were expressed as percent values; however, the underlying changes over time are the same as in the past WSP integrated status assessments.

- *Non-Cyclic Fraser Sockeye CUs; capturing uncertainty in biological benchmark estimates:* This is specific to non-cyclic Fraser sockeye CUs. Uncertainty in *relative abundance* metric benchmarks was incorporated into data summaries in two ways to support the WSP integrated status assessments (Grant and Pestal 2012; Grant et al. 2020): a) *Relative abundance* benchmarks and associated metric statuses were presented across a range of probability levels (10% to 90%) to incorporate assessment uncertainty in the benchmark estimates; b) benchmarks, and associated *relative abundance* statuses, were also shown using multiple stock-recruitment models for each CU, where appropriate.

  To streamline the WSP rapid status approach, we chose one value for each upper and lower benchmark per CU. We used the 50% probability level estimates of the Ricker model-derived benchmarks for non-cyclic Fraser sockeye CUs, as these benchmarks held the most weight in the completed Fraser sockeye WSP status assessments. Fraser sockeye cyclic CUs are described in a subsequent bullet. Note that assessment uncertainty in *relative abundance* benchmark estimates was also presented in the Southern BC Chinook and Interior Fraser coho WSP integrated status assessments; however, *relative abundance* metric statuses were not explicitly presented across benchmark probability levels for these assessments. Future work could test the sensitivity of WSP rapid statuses to alternative probability levels for the benchmark values in the *relative abundance* metric.

- *Highly cyclic Fraser Sockeye CUs*: In the current assessment, five of the 24 Fraser sockeye CUs are considered cyclic: Takla-Trembleur-EStu; Shuswap-ES; Quesnel-S; Takla-Trembleur-Stuart-S; Shuswap-L. Cyclic CUs exhibit persistent abundance patterns that include one large and three smaller abundance years over a four year period, which represents a cycle period of predominantly four year old Fraser sockeye. The larger abundance year is referred to as the 'dominant' cycle year, and the smaller abundance years are referred to as 'off-cycle' years. These highly cyclic Fraser sockeye CUs presented unique considerations in the integrated status assessments (Grant et al. 2011; Grant & Pestal 2012; Grant et al. 2020). We calculated metrics for these cyclic Fraser sockeye CU cases as follows:

  - *Relative abundance metric*: Data summaries used in the first WSP Fraser sockeye integrated status assessment did not present *relative abundance* metrics for cyclic Fraser sockeye CUs (Grant & Pestal 2012). In the second WSP Fraser sockeye integrated status assessment, Larkin model-derived *relative abundance* benchmarks were produced for each of the four cycle-lines of each cyclic CU, across a range of probability levels (Grant et al. 2020). Abundances from each recent cycle line year were compared to corresponding cycle-line benchmarks to determine statuses for each of the four cycles on this metric. We used the 'dominant' cycle spawner abundance and the 50% probability level benchmark in the *learning data set.* This was the rationale provided by experts to designate WSP integrated statuses by experts in the second integrated status assessment (Appendix B; Grant et al. 2020).

  - *Absolute abundance metric*: Generally, during the WSP integrated status assessment workshops, the experts considered the distribution of recent abundances (annual values) against the *relative* and *absolute abundance* benchmarks when determining status. The more years that fell below the lower benchmarks on these metrics, the more likely the CU was *Red*, though experts were less likely to downgrade statuses for cyclic CUs than non-cyclic CUs, particularly when the recent dominant cycle abundance was relatively high. We used the most recent generation geometric average to compare against the *absolute abundance* benchmarks (1,000 and 10,000) in the *learning data set*. This was identical to the approach used for non-cyclic Fraser sockeye CUs.

- *Wild vs. Enhanced:* In the Southern BC Chinook WSP integrated status assessment (DFO 2016), experts completed three status assessments for each case. They evaluated the CU (i.e. fish from wild spawning sites), the enhanced unit (EU; fish from sites with moderate or high enhancement), and the total unit (TU; all sites). We include only the metrics and associated status for the CU (wild) to ensure consistency across species.

2.2.4.2 Retrospective (Out-of-Samples) Data Set Metrics

For the *retrospective (out-of-samples) data sets,* we included the same suite of metrics for each Fraser sockeye and Interior Fraser coho CU that was used in the WSP integrated status assessments. A CU's r*elative abundance* metric benchmark stayed the same throughout the *retrospective (out-of-samples) data set.*

Specifically, for Fraser sockeye CUs, we used the benchmark estimates from the 2017 re-assessment (Grant et al. 2020), not the benchmark estimates from the first assessment in 2012 (Grant & Pestal 2012). This included Ricker model derived benchmarks for non-cyclic CUs, and Larkin model derived benchmarks for cyclic CUs on the dominant cycle line.

For Southern BC Chinook, however, more recent work was used to identify the appropriate metrics for the *retrospective (out-of-samples) data set*, the results of which deviate from those used in the WSP status assessments, and *learning data set,* for some CUs (Brown et al. 2020). We used the metric usability categorizations from Table 5 in Brown et al. 2020, and the most recent site classifications, to define which CUs and metrics to include in the *retrospective data set*. The classification of enhancement level was also updated for some sites in some CUs since the WSP integrated status assessments.

- Data treatment for the three WCVI CUs was revised for consistency with a recent case study reviewed by CSAS (Holt et al. 2023b) and further revised based on guidance from CU experts. Specifically, the set of indicator sites used to build the CU-level time series was revised based on new information (PNI, Proportionate Natural Influence, Withler et al. 2018) on hatchery contribution. This PNI-based revision to CU-level time series is potentially applicable to many other SBC Chinook CUs that are currently data deficient due to the enhancement rating for indicator sites, and could greatly expand coverage of the WSP rapid status scan (e.g., Inner South Coast).
- New data were provided by BC Interior staff for Okanagan Chinook, using estimates of natural-origin spawners developed by the Okanagan Nation Alliance.

For the retrospective analysis, and for future case studies, we standardized the metric calculations. The *WSPMetrics* R package (Pestal and Holt 2020) implements the same approach used in the *learning data set*, producing a single value for each metric as inputs to the algorithms.

2.3 PERFORMANCE EVALUATION OF WSP RAPID STATUS ALGORITHMS

2.3.1 General Approach

To evaluate the performance of candidate algorithms, we first established agreed-upon criteria to identify which aspects of model performance to prioritize. These criteria determined which performance measures and data to use, how to weight performance measure results, and what other criteria to consider when evaluating algorithms.

Performance criteria were used to both define and evaluate algorithms:

1. They were first used for pre-screening algorithm variations that were explored in the development step, leading to a shortlist of candidate algorithms.
2. Then they were used for a detailed performance evaluation of the candidate algorithms.

Depending on the criterion, the algorithm performance evaluation was either qualitative or quantitative. Quantitative evaluations were done by using the *learning data set* to compare predicted values (in this case: WSP rapid statuses) to observed values (in this case: WSP integrated statuses from expert consensus) and quantify the magnitude and direction of errors.

We next performed two sensitivity tests to (1) evaluate the stability of the algorithm statuses over time (*retrospective (out-of-samples)* test), and (2) test how the inclusion or exclusion of the *relative abundance metric* affects WSP rapid statuses produced by the different algorithms. Status is intended to trend over time without large annual variations. The first sensitivity analysis evaluates the extent to which this holds for the candidate algorithms, using the *retrospective data set*. The second sensitivity analysis evaluates the extent to which the accuracy of candidate algorithms depends on the availability of the *relative abundance* metric, which typically is only available for data-rich CUs. For this test we excluded the *relative abundance* metric from the *learning data set* and evaluated algorithm performance.

The rest of this section documents the criteria, performance measures, and qualitative considerations, including sensitivity tests, we used.

2.3.2 Criteria for Selecting WSP Rapid Status Algorithms

The criteria we used to guide the construction, evaluation, and selection of candidate algorithms are as follows:

1. Algorithms should have relatively low error rates when comparing WSP rapid statuses to integrated statuses, the latter which are assumed to be ‘true’ statuses.

2. Algorithm errors should be precautionary, meaning that estimated rapid statuses should err on the side of being poorer, indicating a higher risk of extirpation, when compared to ‘true’ integrated WSP statuses. For example, if a ‘true’ integrated WSP status is *Amber,* a rapid status error should be more likely to be *Red* over *Green*.

3. Algorithms must be broadly applicable across CUs with different data types and metric availability.

4. Algorithms that estimate WSP rapid status for three main status zones: *Red*, *Amber*, and *Green* are preferred.
5. Algorithms should reflect thresholds that emerged from those distinguishing statuses in WSP integrated status assessment. These tend to be equal to or more biologically conservative than WSP benchmarks for individual metrics from Holt et al. (2009).
6. Algorithm decisions should adhere to the logic applied in the WSP integrated status assessments. This includes following common rationale applied in the detailed WSP status assessment processes, as documented in the CU status narratives reprinted in Appendix B, which includes extracts from Grant and Pestal (2012), Grant et al. (2020), Brown et al. (2014) and Parken et al. (2014).

Performance of the algorithms on criteria 1, 2, and 3 can be quantified using error rates, measures of bias (specifically over-prediction), and completion rates, respectively. Completion rate is the proportion of the 65 cases in the *learning data set* for which WSP rapid status could be assigned. These quantitative performance measures were calculated across all CUs and by species.

Performance on criterion 4 can be easily evaluated by checking that all three simple WSP integrated status zones (*Red*, *Amber*, and *Green*) are included as branches of the algorithm trees.

Performance on the remaining criteria, 5 and 6, is subjective and was evaluated by expert opinion.

We iteratively evaluated and altered candidate algorithms based on their performance against these criteria. For some algorithms, this was done by adjusting the CART tree fit settings. For other algorithms we actively revised or reorganized the decision nodes. Section 2.4 describes the details.

2.3.3 Quantitative Performance Measures

For the 65 cases in the *learning data set* (Section 2.2), we compared WSP rapid statuses generated by each of the candidate algorithms to the existing WSP integrated statuses (considered ‘true’ statuses).

We used the entire *learning data set* to evaluate performance of candidate algorithms using the six criteria above. Due to the small sample size, we did not use cross-validation approaches that split these data into *learning* and *testing data sets,* as is commonly done for forecasting models (see review in MacDonald and Grant 2012). Cross-validation is generally recommended to minimize the risk of over-tuning models to the idiosyncrasies of the data being used; this is intended to minimize overly optimistic expectations for how models will perform with new data sets (Picard and Cook 1984). However, the *learning data set* had a relatively small sample size, making cross-validation inappropriate (Picard and Cook 1984).

Instead, to prevent overfitting the candidate algorithms to CUs and years in the *learning data set*, and to ensure the algorithms were broadly applicable to all BC and Yukon CUs, we applied the following methods:

- We developed performance criteria (Section 2.3.2) to guide the construction, evaluation and selection of candidate algorithms. If we had relied exclusively on model performance determined through cross-validation, this would have increased the risk of selecting an algorithm that is ‘overfitted’ to the *learning data set.*

- Algorithm development included both *fitted* and *constructed* algorithms:
  - Three *fitted* algorithms were based on CART analyses (Section 2.4.2), which uses cross-validation to determine error rates and types. Using CART analyses, algorithm fit is determined by balancing error rates and types, and tree complexity. Different fitted algorithms were developed by altering both the complexity setting from low to high, and altering the use of metric values or statuses. CART analysis was conducted using the R package *rpart* (Therneau and Atkinson 2019).
  - Four *constructed* algorithms (Section 2.4.3) were developed using the CART algorithms as a baseline. These constructed algorithms were built to more closely align with the algorithm criteria in Section 2.3.2. and incorporate common rationale extracted from existing WSP integrated status assessments. Considering common rationale that would be applicable to a broad range of CU data types reduced the risk of overfitting algorithms to the *learning data set.*
- We conducted a *retrospective* (*out-of-sample*) test with the seven candidate algorithms for years that do not have WSP integrated status assessments completed. Another companion report conducts an out-of-sample test for CUs not previously assessed (Pestal et al. 2023). Experts verified status results in these cases to confirm the applicability of the algorithm(s) to these new assessed years.

To calculate prediction errors we first converted statuses to scores from 1 = *Green* to 5 = *Red* (Table 4). We then calculated the difference between WSP integrated status scores and the WSP rapid status scores (i.e. observed-predicted). A negative error means that the algorithm predicted a poorer status than the WSP integrated status. Note that candidate algorithms differed in terms of possible outcomes (e.g. whether *Red/Amber* and *Amber/Green* options are included), and the status scores were adapted accordingly. Section 2.4.4 describes the status scales (Table 4) and how they were used.

We used the following quantitative performance measures to compare algorithm performance for all cases:

- *Number and percent correct:* the total number of cases and the percent of cases where the WSP rapid status matches the WSP integrated status (‘true’ status). This measures alignment with Criterion 1. Note that percent correct is calculated from the number of completed cases (see below), not the total number of cases (e.g. if an algorithm can assess 40 of the 65 cases, and 30 WSP rapid statuses match the WSP integrated statuses, then the percent correct is 30/40 = 75%, not 30/65 = 46%).
- *Number and percent over-predicted:* the total number and percentage of cases with positive errors in status estimates, where the WSP rapid status assigned by the algorithm is better than the WSP integrated status. This measures alignment with Criterion 2. Percent over-predicted is calculated from the number of completed cases.
- *Number and percent completed:* the number and percent of cases where the algorithm was able to generate WSP rapid statuses. This partially measures alignment with Criterion 3 to the extent that different data types and metric availability are represented in the *learning data set*.
- *Median, Mean, and Range of Prediction Errors:* summary statistics that describe the distribution of prediction errors and identify any bias.

In addition, we cross-tabulated WSP integrated statuses against the WSP rapid statuses predicted by an algorithm. The frequency of each type of possible error resulting from misclassification was estimated. For example, a CU with a *Green* WSP integrated status that is misclassified by the algorithm as *Amber* will have the same error of +2 as an *Amber* CU misclassified as *Red*, but the biological implications of the error are different. We present the details for each error type to explore these differences. The practical implications of observed errors were evaluated qualitatively through discussions with species experts (C. Parken and A. Vélez-Espino, DFO, pers. Comm).

2.3.4 Sensitivity Test 1: Retrospective (Out-of-Sample) Test

For the *retrospective (out-of-samples)* test, we applied all candidate algorithms to the data and metrics for each year in the *retrospective data set*, which does not include any 'true' WSP integrated statuses. Therefore, the *retrospective (out-of-samples)* test can only evaluate status changes over time.

We looked at whether the WSP rapid statuses for a CU were stable or gradually changed over time (desired properties), or bounced between status zones frequently between years, possibly due to assessment errors. Status is intended to focus on the overall signal in the data and not vary inter-annually in response to noise.

We also compared the time series of annual WSP rapid statuses between algorithms to look for similarities and differences in these patterns. For example, several algorithms may identify a worsening status for the same year, even if they assign different statuses, resulting in similar patterns.

The *out-of-sample* test also serves as a practical test of how the selected algorithm can be used going forward, with annual status updates and decision-support tools focusing on status changes over time.

2.3.5 Sensitivity Test 2: Excluding Relative Abundance Metrics

The *relative abundance* metric is not available for all CUs (see Section 2.2). In fact, most Pacific salmon CUs likely will not have this metric. Therefore, for CUs in the *learning data set* with *relative abundance* metrics, we tested the influence of including or excluding this metric on the WSP rapid statuses generated relative to the 'true' WSP integrated statuses.

There were 37, out of 65 cases in the *learning data set,* that have values for the *relative abundance* metric. We compared the inclusion or exclusion of these values for the *relative abundance* metric for these cases, on the WSP rapid statuses, compared to 'true' WSP integrated statuses.

2.4 DEVELOPING A SHORTLIST OF CANDIDATE ALGORITHMS

2.4.1 Overview

We developed two types of algorithms (Appendix D):

- *Fitted*: based on Classification and Regression Tree (CART) analyses;
- *Constructed*: candidate CART trees adapted based on common rationales extracted from existing WSP integrated status assessments.

We explored algorithms that predict WSP rapid statuses based on either the metric values (e.g. *percent change* shows a decline steeper than -25%) or the metric statuses (e.g. *percent change* is *Red*). Similarly, we explored algorithms that predict all five WSP integrated status zones: *Red*, *Red*/*Amber*, *Amber*, *Amber*/*Green*, and *Green*; just the three simplified status zones: *Red*, *Amber*, and *Green*; or two simplified zones: *Red*, and *Not Red*. These align with, respectively, the 5 status scale, the 3 status scale and 2 status scale for error calculations described below (Table 4; Section 2.4.4).

2.4.2 Fitted Algorithms using CART Analysis

Classification and Regression Tree (CART) analyses are widely used in decision analysis and machine learning to identify complex patterns in data and develop algorithms for classification of new cases (Ripley 1996). The approach is very versatile but comes with highly specialized terminology (Table 2).

Briefly, CART searches for a binary split in available data or cases, which uses a criterion to divide the original group of cases into two smaller groups of cases. Tree branches are added as these new groups are further split into even smaller groups. In technical terminology, the approach uses machine learning methods to build a dichotomous key to the data.

What determines the 'best' grouping of cases depends on error rates (i.e. number of incorrect classifications), error type (e.g. false positives vs. false negatives in a classification tree that screens for a medical condition), and tree complexity (i.e. the number of branches on the tree). In CART, the fitting step balances the number of branches (complexity) against the magnitude and type of misclassifications.

The strength of binary trees comes from *recursive partitioning*. At each node, a single criterion is used to split the cases into two more homogeneous subsamples. That means that a binary tree can pick up conditional interactions between variables in a very straight-forward way. As an illustration, consider a field guide for species identification. Once the data on species traits have been worked through several of the steps and the options are narrowed down to two possible species, then a single easily identifiable traits can tell them apart. However, that same characteristic would not yield a proper classification if it were applied to the entire sample of possible cases.

We used the R package *rpart* (Therneau and Atkinson 2019) to fit classification trees to our *learning data set* of 65 completed WSP status assessments, using the metrics as predictor variables and WSP integrated statuses as the response variables. *Rpart* uses cross-validation to estimate error between predictor and response variables. We explored many alternative fitted trees working through variations of response variables, predictor variables, model fits, and data subsetting (Table 3).

From the many variations of fitted algorithms that we explored, we selected a shortlist of candidate algorithms. Shortlisted CART algorithms were chosen to bookend the range of complexity possible through this analysis, including one very simple algorithm (*Minimalist*) and one very complex (*Fancy Pants*) (Table 6). This pre-screening step used the criteria, quantitative performance measures, and qualitative considerations documented in Sections 2.3.2 and 2.3.3, which are then used again to evaluate the performance of the shortlisted algorithms.

We identified three candidate *fitted* algorithms for detailed performance evaluation (Table 6):

Minimalist

The *Minimalist* algorithm (Appendix D.1) was created by setting the complexity penalty high and working with the simplified status categories (*Red, Amber, Green*). It is the simplest model fit using the CART analysis (Table 6). The intent with this algorithm is to have the greatest applicability to the broadest range of CUs. This algorithm uses only trend-based metrics (*long-term trend* and *percent change*), as these metrics are the most likely to be available across the range of Pacific salmon CUs in the Pacific Region.

The *Minimalist* algorithm relies on metric values, not metric statuses, as predictor variables. Therefore, the splitting thresholds are extracted by the CART fit from the expert assessments, and do not simply carry over the official metric benchmarks.

Fancy Pants

The *Fancy Pants* algorithm (Appendix D.2) was created by setting the complexity penalty low and working with the full range of status categories (including *Red/Amber* and *Amber/Green*). *Fancy Pants* is the most complex algorithm fit by the CART analysis (Table 6). This algorithm uses all of the available metrics and is the only algorithm that can assign statuses to all five status zones: *Red, Red/Amber*, *Amber, Amber/Green* and *Green.*

Categorical Realist

The *Categorical Realist* algorithm (Appendix D.3) was included as a candidate algorithm to specifically incorporate absolute abundance vs. relative index data types into the CART analysis binary splits. The *Categorical Realist* algorithm was developed specifically to address algorithm Criterion 3 and 5 (Table 6). Therefore, this algorithm uses metric statuses, instead of values, to separate tree branches, since metric statuses are determined using the WSP benchmarks (Holt 2009). We incorporated an initial step that separates the data based on data type: absolute abundance, or relative index. Trees were fit individually to each data type, then combined into one tree. The drawback of this algorithm is that it only assigns *Amber* and *Red* statuses, there is no branch for the *Green* status zone (Criterion 4). This is a result of how the fitted tree splits the sample and assigns a designation to each subset. In this particular case none of the four endpoints ('leafs') has a majority of samples with *Green* status. However, it is extremely simple, relying on only two metrics once the initial data type screening is complete.

2.4.3 Constructed Algorithms

Using the CART-fitted algorithms as a baseline, we built four constructed algorithms to more closely align with the six performance criteria identified in Section 2.3.2. One algorithm is relatively simple, aptly named *Simply Red;* and three are more complex, named *Learning*

Tree 1, 2 and 3 (Table 6; Appendix D). The more complex *Learning Tree* algorithm represents an evolution of improvements over each subsequent version, representing the adaptive nature of this algorithm. This algorithm illustrates the approach we are proposing for future implementation, as new WSP integrated status assessments are completed for more CUs and Areas, which may require further improvements to the WSP rapid status algorithm.

Simply Red

The *Simply Red* algorithm (Appendix D.4) was designed to specifically address Criterion 2. This criterion calls for algorithms to produce WSP rapid statuses that are more biologically conservative (i.e. err on the side of poorer status) than WSP integrated statuses (Table 6). To do this, we mined CART-fitted models to identify nodes where *Red* statuses were assigned, and combined these into one tree that includes all criteria used to flag *Red* CUs. The algorithm uses metric values identified in the *Minimalist* and *Fancy Pants* fitted algorithms to identify binary splits. However, these values have been adjusted to align with WSP benchmarks and COSEWIC criteria, where appropriate (Criterion 5).

Since the objective with this algorithm is to identify *Red* CUs, it assigns only two statuses: *Red*, and *Not Red* (Criterion 4). The algorithm does not assign the *Amber* and *Green* status zones, therefore, its applicability for WSP rapid status assessments is limited to scanning for poor status CUs.

The Learning Tree family of algorithms

For the *Learning Tree* family of algorithms (Appendices D.5, D.6 and D.7; Table 6), our objective was to follow all of the criteria as closely as possible, though we placed the greatest emphasis on two qualitative criteria:

- improving algorithm applicability to populations with different metric availability (Criterion 5), and
- algorithm adherence to the logic applied in the WSP status integration processes (Criterion 6).

This was to ensure that the algorithm can be applied to CUs outside of those included in the *learning data set*. The intention behind this series of algorithms is that they will continuously improve as more WSP integrated statuses across more CUs become available.

First, we examined documentation of the completed WSP integrated status assessments in published peer-reviewed reports (Grant and Pestal 2012; Grant et al. 2020), and unpublished reports (Brown et al. 2014; Parken et al. 2014) (See Appendix B). We considered both the WSP integrated statuses and the narratives that experts developed when they assigned CU status (see Appendix B). Using this information, we extracted common rationale, either identified explicitly for the group of CUs being assessed, or indirectly through repeated mention across CUs. From these common rationales we identified those that were most broadly used across all species, and those that pertained to specific data types and situations.

Next, we reviewed the CART-fitted trees for common decisions made in these analyses, and compared these to the common rationale from the WSP integrated status assessments. From this, we identified essential tree elements to include in the constructed algorithm from the existing branches of the fitted trees. We then incorporated important WSP integrated status assessment considerations that were missing from the CART-fitted trees where we could. Constructing this algorithm by hand allowed us to include conditional rules within tree branches that can better capture the nuances of the decision-making processes of the WSP integrated status assessments.

Development of the *Learning Tree* algorithms was highly iterative. As each branch evolved, this algorithm was evaluated for biological rationale and consistency with the WSP integrated status assessments, and error rates (Section 2.4.4) were investigated. Where there were differences between algorithm-generated WSP rapid statuses and the WSP integrated statuses, we had a deeper look at the documented rationale behind the WSP integrated status to identify missing components, and/or alternative metric breakpoints, as described below. These considerations were iteratively added to the algorithm where possible, hence the *Learning Tree* has three alternative versions so far (*Learning Tree 1, 2 and 3*) (Appendix D.5, D.6 and D.7). *Learning Tree* 3 is the most recent iteration of this algorithm (Appendix D.7).

The initial *Learning Tree* development used components of the *Minimalist*, *Fancy Pants*, and *Categorical Realist* algorithms, guided by common rationales used to assign status from past WSP integrated status assessments. Common rationales details are provided in Appendix B, and summarized below:

- Where *relative abundance* metrics are available, these highly influence a CU's status, independent of other metrics in many cases.
- Where absolute abundance data exist, *absolute abundance* metrics should be scanned against COSEWIC criterion D1 on small population sizes (>1,000 fish). This step is not included in any of the fitted algorithms. In adding this step, we applied a precautionary buffer to the COSEWIC threshold for small populations (1,000 fish), setting the threshold at 1,500 in the *Learning Tree 3.* This was to account for how this metric was treated by experts in the workshops, where CU statuses may be downgraded if one year in a generation fell below 1,000, if the estimates were considered uncertain, or if the generational average was close to the 1,000 threshold.
- Similar to the precautionary buffer on the lower benchmark of the *absolute abundance* metrics (previous bullet), we also added a 10% buffer to the upper threshold for the *relative abundance* metric. The standard WSP metric uses 80% of S_{msy} as the upper benchmark for *relative abundance*, but the *Learning Tree* algorithms treat abundances that are only a little bit above the benchmark as a flag for *Amber* status.
- *Long term trends* and *percent change* are most heavily relied upon only in the absence of other metrics, and where COSEWIC criterion D1 is nottriggered

Using these WSP common rationale as the base for tree structure, we pulled nodes from the existing CART-fitted trees to build two versions of the *Learning Tree*.

Learning Tree 1 was initially constructed using metric value thresholds from the *Minimalist* and *Fancy Pants* algorithms, or values based on the WSP benchmarks. Some thresholds were then adjusted to better align with patterns in the data and ensure that they are precautionary (Criterion 2). For example, the *percent change* threshold (i.e. short-term trend over 3 generations) was changed from -80% in the *Minimalist* to -70% in the *Learning Tree* algorithms. With this change, a less steep decline is needed to trigger the criterion.

Learning Tree 2 uses metric statuses inferred from the threshold values extracted from the CART analyses, or from the WSP assessments. *Learning Tree 2* was produced to ensure that Criterion 5 was being met, and metric thresholds are biologically justifiable.

The resulting *Learning Tree* algorithms are applicable across data types and metric availability (Criterion 3). In contrast to the other candidate algorithms, *Learning Tree 1* and *2* use all available metrics: they use both the *relative abundance* and *absolute abundance* metrics where applicable, but also provide status pathways for CUs where these metrics cannot be calculated.

Learning Tree 3 evolved from *Learning Trees 1* and *2* after reviewing results for Southern BC Chinook and Interior Fraser coho with experts on these species and their WSP integrated status assessments. *Learning Tree 3* includes the following considerations to improve results for Interior Fraser coho in particular; however, these additions are likely to be widely applicable across all Pacific salmon CUs:

- CUs are scanned against the population size threshold used as part of COSEWIC Criterion C, where available. This corresponds to the upper benchmark used in the *absolute abundance* metric. If abundance falls below this threshold (10,000 fish), the only status options are *Red* and *Amber*, because COSEWIC would be more likely to assign a *Threatened* status when this criterion is met (Section 2.2.2).
- The order of branches was changed to check whether the WSP rapid statuses were sensitive to changes in the sequence of decision nodes and was then settled to complete COSEWIC scans prior to other steps. Specifically, the first step in *Learning Tree 3* is to check whether absolute abundance data are available, and if so, whether they fall close to or below the COSEWIC thresholds for small populations.
- The consideration of *long-term trend* and *percent change* (i.e. recent short-term trend) was fine tuned in *Learning Tree 3*. That part of the tree applies to cases where *relative abundance* metric is not available. Specifically:
    - The node identified by steep recent decline (*percent change* < -70%) was changed from *Amber* in *Learning Tree 1* to *Red* in *Learning Tree 3* to be consistent with COSEWIC criterion A, which states "*An observed, estimated, inferred or suspected reduction in total number of mature individuals over the last 10 years or 3 generations, whichever is the longer, where the causes of the reduction are: clearly reversible and understood and ceased, based on (and specifying) any of the following: For endangered -70%, for Threatened -50%. If the causes of decline are NOT known and reversible, this % is -50% for endangered and -30% for threatened.*" So any CU that has -70% decline will be *Red* by COSEWIC regardless of abundance, and *Learning Tree 3* was adapted to be consistent with that criterion.
    - After this change, that part of *Learning Tree 3* needed another step to split out *Green* vs. *Amber*, so the corresponding criterion from the fitted *Minimalist* algorithm was included as the final step.

2.4.4 Error Calculations on Alternative Status Scales

For the completed status assessments in the *learning data set*, we compared WSP integrated statuses to the WSP rapid statuses to assess status classification error. To do this we converted all statuses to numeric scores (Table 4). WSP integrated statuses were originally designated for five status zones (*Red, Red/Amber*, *Amber*, *Amber/Green* and *Green*). These were converted to scores on three different status scales to match the status scales used by different WSP rapid status algorithms. WSP rapid status algorithm statuses were also converted to scores from 1 to 5 on their respective status scales (Table 4 and Table 6).

The *5 status scale* directly aligned to the five status zones of the WSP integrated status assessments. WSP integrated statuses converted to a 5 point scale includes: *Green* = 1; *Amber/Green*=2; *Amber*=3; *Red/Amber*=4 and *Red*=5. This scale is most appropriate for the

Fancy Pants algorithm, which also provides status for all 5 status zones (Table 6). If *Fancy Pants* assigns a *Red* status (score 5) on its 5 point scale, and the WSP integrated status is *Amber* status (score 3) on its 5 point scale, then the error is 2. However, other algorithms were also compared to the 5 point WSP integrated status scale using their scores on the scale that aligns with their results (Table 4)**.** For example, if the WSP integrated status was *Amber/Green* (score = 2) and the *Simply Red* status was *Red* (score = 5 on the 2 *status scale),* then the error is 2-5 = -3.

The *3 status scale* was applicable to WSP rapid status algorithms that assign only simplified statuses (*Red=5, Amber=3, Green=1*) and do not include mixed status zones (*Red/Amber* and *Amber/Green*) (Table 4). To convert the mixed integrated statuses to simplified status zones we used only the lower zone from the mixed status. For example, a CU with an WSP integrated status of *Red/Amber* became *Red* on the 3 status scale (Table 4). This was necessary to accurately represent error across algorithms, without over-estimating the error rate of those algorithms that produce WSP rapid status using only the *Red*, *Amber*, and *Green* status zones. It weights the algorithms towards being more biologically conservative in their assignment of statuses. This status scale was aligned with the *Minimalist* and *Learning Tree 1,2 and 3* algorithms. The *Categorical Realist* also aligns with the 3 status scale, except that it only produces *Red* or *Amber* statuses, not *Green*. Again, other algorithms were also compared using the 3 point scale to evaluate performance.

The 2 *status scale* was applicable to the *Simply Red* algorithm that assigns only *Red=5* or *Not Red=2* statuses (Table 4). For this scale we converted the WSP integrated status of *Red* or *Red/Amber* into a *Red* status, and all other statuses (*Amber, Amber/Green,* or *Green*) into *Not Red*. This was done for similar rationale as the 3 status scale. Specifically to accurately represent error across algorithms without over-estimating the error rate for the algorithm that was designed using only the *Red* and *Not Red* status zones. Although this status scale was aligned to the *Simply Red* algorithm, other algorithm results were also compared to this 2-point scale.

It is most appropriate to review an algorithm's performance by comparing the algorithm status and the WSP integrated status on identical status scales.

- **5 status scale:** *Fancy Pants*
- **3 status scale:** *Minimalist, Categorical Realist* (although this only produces *Red* and *Amber* statuses), *Learning Trees 1 to 3;*
- **2 status scale:** *Simply Red*

If a mis-matched status scale is used to assess error for an algorithm, it is not entirely an apples-to-apples comparison. For example, if a case's WSP integrated status is converted from *Amber* (score: 3) on the original 5 status scale, to *Not Red* (score: 2) on the 2 status scale, then algorithms with 3 status scales scores (either *Green*=1; *Amber*=3; or *Red*=5) will never align to the 2 status scale score of *Not Red*=2.

We evaluated performance of each algorithm on all three status scales, in order to allow comparisons across algorithms, but highlight for each algorithm the scale that is most appropriate for evaluating performance of that algorithm on its own.

Each status scale is informative for a different purpose. The 5 status scale captures all the nuances of the integrated status assessments from the expert workshops, but can only be compared to the WSP rapid statuses from the *Fancy Pants* algorithm. The 3 status scale matches the outputs for most of the other candidate algorithms, but a direct comparison to integrated statuses on a 5 status scale would distort the error calculations (i.e. an *Amber* designation produced by *Learning Tree 3* for a CU with integrated status *Amber/Green* is not wrong, rather it is correct at a coarser level of resolution). The 2 status scale is the only one that gives an accurate picture of how the *Simply Red* algorithm performs, but it can also be used to compare all the candidate algorithms on the same scale. Finally, which status scale is the most useful depends on the question being asked. If the intent is to track patterns in status on a fine scale, then the 5 status or 3 status scales are more useful. If, however, the intent is to identify the number of *Red* CUs, then algorithm performance on the 2 status scale is highly informative.

2.5 CAPTURING CONFIDENCE IN WSP RAPID STATUS DESIGNATIONS

Confidence in the WSP rapid status designations can be increased through careful screening of the data and metrics being used (i.e. quality control of inputs). Once the inputs have been vetted, confidence in the status designations can be quantified.

Robustness of fitted trees based on CART models can be evaluated by comparing alternative algorithms, but in our analysis we are already identifying the single best performing algorithm based on the set of criteria presented earlier.

During the WSP integrated status assessment processes, confidence in status assessments increased when there was convergence in statuses across individual metrics (i.e., cases where the *absolute abundance, relative abundance, percent change* and *long-term trends* all indicated the same status). However, it is not clear how agreement across metrics could be considered consistently when there are mixed signals across metrics, due to the many possible variations.

We therefore capture confidence in the WSP rapid status results based on the type of information used to assign status, which has two components:

1. Data screening and metric applicability (Section 2.5.1) and;
2. Assigning confidence based on algorithm node using the branches of the recommended algorithm, which is based on the type of data available and the sequence of criteria used to assign the status (Section 2.5.2).

2.5.1 <u>Data Screening and Metric Applicability</u>

To develop the *learning data set* used in the WSP rapid status approach, we used the data, specifications and metrics identified in past WSP integrated assessment processes (Grant et al. 2011, 2020; Grant and Pestal 2012; DFO 2015, 2016). These processes relied on years of work selecting and treating the data, and identifying relevant metrics, benchmarks, and specifications (such as average generation length) required to assess WSP status. Processes also relied on workshops and CSAS peer review to finalize data, metrics, metric interpretation, and status assignments.

Through these past processes, DFO Area stock assessment staff compiled the data sets used in this WSP rapid status assessment work. This data set was developed through their understanding of DFO field survey and escapement estimation approaches, and also through collaborations and engagement with external projects led by First Nations, consultants, etc. Through their local expertise on particular CUs, and their engagement with external local experts, data applicable for use in WSP status assessments is identified.

Poor quality data are not included in the data sets used for WSP status assessments as part of the data processing step. Area staff are involved in determining which data are included in the CU data sets, and which years require infilling for metric calculations. This step includes collaborative work with externals who are leading or collaborating on relevant stock assessment projects.

DFO Area staff are also involved in selecting which metrics can and cannot be calculated from the data sets they provide. This step pre-screens all poor-quality data out of the process and ensures that the data are appropriate for the metrics being calculated. For example, *absolute abundance* metrics are only calculated where absolute abundance data are present or otherwise deemed appropriate. As a further screen, apart from *absolute abundance*, each metric either requires that the time-series provided be of a certain length in order to calculate the metric (e.g. need 3 generations of data to calculate *percent change*) or requires benchmarks provided by Area staff, which generally require adequate stock-recruit data.

2.5.2 Assigning Confidence Based on Algorithm Node

Greater confidence in status is associated with particular metrics and status results. In the WSP integrated status assessments, assigned statuses were more consistent across experts for some cases than others. In particular, cases that had *absolute* and *relative abundance* metrics were more consistently assessed than those with only trend metrics and relative index data (Figure 3 and Figure 4). To apply these metrics, a CU must have higher quality data to be able to estimate benchmarks, or estimate this metric's annual value for comparison with its benchmarks. Therefore, we have more confidence in statuses that are assigned using the *absolute* and *relative abundance* metrics, than statuses assigned using *long-term trend* metrics.

The *long-term trend* metric compares a CU's metric value (ratio of the current generational average spawner abundance to the long-term average) to the metric's benchmarks. The CU's value for this metric is influenced by the length of the time series and degree of fisheries exploitation that occurred early in the time series. This metric can also be calculated for lower quality data, including indices of abundances. For these reasons, this metric was considered less reliable to assess status by experts in past WSP integrated status assessment processes.

To account for these differences in confidence identified from past WSP integrated status assessment processes, we used the branches of the algorithm itself to identify confidence in the statuses being assigned, based on the combination of metrics, metric values, and data types that determine each status node. Through expert judgement we can bin the end nodes into three confidence zones: *High, Medium,* or *Low*, and then evaluate this binning by referring to the *learning data set* CUs that end up in each zone.

Confidence ratings below were applied to each end node of the *Learning Tree 3* algorithm as follows (Figure 20):

- *High confidence - Red*: either *absolute abundance* is available and falls below 1.5 times the lower benchmark on this metric **(node 3)**, **OR** *relative abundance* benchmarks are available and generational average spawner abundance falls below the lower benchmark **(nodes 19 or 23)**.

- *High confidence – Green:* abundance is above the upper benchmark on the *absolute abundance* metric, or this cannot be assessed **AND** *relative abundance* benchmarks are available and generational average spawner abundance falls above 1.1 times the upper benchmark **(node 36)**.

- *High confidence – Amber:* abundance is above the upper benchmark on the *absolute abundance* metric, or this cannot be assessed; *relative abundance* benchmarks are available and generational average spawner abundance fall between the lower and 1.1 times the upper benchmarks **(node 37)**.
- *Medium confidence – Red:* (1) *absolute abundance* falls between the upper and 1.5 times the lower benchmarks and status is based on *long-term trend* **(node 21)**; or, (2) abundance is above the upper benchmark on the *absolute abundance* metric, or cannot be assessed, *relative abundance* metrics are not available but status can be assessed based on *long-term trends* alone **(nodes 17)** or with both *long-term trend* and *percent change* (**node 33)**.
- *Medium confidence – Amber*: either (1) have *relative abundance* benchmark and *absolute abundance* is between the upper and 1.5 time the lower benchmark **(node 22)**, (2) *relative abundance* metrics are not available, but *absolute abundance* is between the upper benchmark and 1.5 times the lower benchmark, and based on *long-term trend* (**node 20**)
- *Low confidence - Amber*: abundance falls above the upper benchmark on the *absolute abundance* metric, or cannot be evaluated on this metric, and *relative abundance* metrics are not available so CU status is assessed based on *long-term trend* and *percent change* **(node 65)**.
- *Low confidence – Green:* abundance falls above the upper benchmark on the *absolute abundance* metric, or cannot be evaluated on this metric; *relative abundance* metrics are not available; status is based on trends alone (*long-term trend* and *percent change*) (**node 64**).

2.6 IMPLEMENTATION OF CANDIDATE ALGORITHMS

Computational implementation of the candidate algorithms required substantial programming. When fitting CART models with the {*rpart*} package in R, the resulting tree object is fully integrated with the generic R functions for working with fitted models. For example, the output includes error summaries like the confusion matrix and surrogate splits for handling missing data (Table 2). The fitted tree can also be applied easily to new data with the *R* function *predict()*. This is extremely efficient when dealing with large data sets and many alternative complex trees, because it automates the full sequence from testing alternative fitting criteria to evaluating predictions.

However, in our particular decision setting we are dealing with a small data set of cases where a broader planning process needs to be able to understand the rationale for each classification (i.e. fully transparent). The surrogate splits were creating challenges during our review of candidate algorithms, and there is no simple option for turning off surrogate splits in the R function, *rpart*().

In addition, there is no simple way for generating tree objects for constructed algorithms in a way that mimics the output from *rpart()* so that it can be fed into *predict*(). We therefore built a custom function, *rapid_status()*, which applies the fundamental logic of the trees generated by *rpart(),* such as the node numbering system described in Table 2, but with hardwired classification steps for each candidate algorithm and customized outputs specifically for our data structure (e.g. error calculations).

The *rapid_status()* function is available as source code and in an R package in a github repository available upon request (sue.grant@dfo-mpo.gc.ca).The code required a lot of trial and error to identify and handle special combinations of metrics values and missing data, so we recommend that any future applications work with the latest version of the function available through this repository, rather than trying to implement themselves the steps in the decision tree diagrams in Appendix D.

3 RESULTS

3.1 PERFORMANCE WITH LEARNING DATA SET

Despite differences among the algorithms and across CUs in terms of available metrics, all algorithms were able to complete WSP rapid status assessments for most of the cases, with number complete ranging from 54/65 (83%) to 65/65 (100%) of the cases in the *learning data set* (Table 7). Most algorithms were also able to achieve a high number of correct WSP rapid status assignments on their associated status scales (see grey highlighted cells in Table 7; Figure 8). Comparisons of WSP rapid statuses generated by each algorithm compared to the WSP integrated statuses are presented for Fraser sockeye (Figures 5 & 6); Southern BC Chinook and Interior Fraser coho (Figure 7).

We focused on the total number of correct cases for comparing algorithms (Table 7; Figure 8), because this performance measure captures both the completion rate and whether the completed assignments are correct. For example, Learning Tree 3 could assign statuses for all 65 cases, and 54 of the assignments were correct (since all cases were completed, this represents 83% of the total 65 cases). By comparison, Fancy Pants completed only 54 cases, but 47 of those were correct (72% of all 65 cases; 87% of completed 54 cases). Depending on which percentage you look at, Fancy Pants performed either much better than Learning Tree 3, or a little bit worse. By focusing on the total number correct, we avoid potential confusion with these alternative percentages.

Patterns in completion rates were similar across species for each algorithm (Table 8 to 10). All *Learning Tree* algorithms had 100% completion rates across species using the full *learning data set*. The *Minimalist* also had close to 100% completion rate across species. Other algorithm completion rates varied across species. The *Categorical Realist* completion rate was lower for Fraser sockeye in particular (78%), but was 100% for Chinook and coho. *Simply Red* and *Fancy Pants* had a lower completion rate for both sockeye and Chinook, but 100% completion for coho.

Most errors across algorithms (>60%) were within one gradation on the 5 status scale away from the WSP integrated status (Figure 9). This was categorized as ‘close’ in Table 7 to 10 (Figure 9).

More detail on errors can be found in Table 13; and Figure 10 to Figure 16. These show the distribution of errors for each algorithm. This provides additional information about the direction and magnitude of the classification errors summarized in Tables 7 to 10.

Appendix E summarizes the performance for each algorithm. Appendix F provides detailed error diagnostics; this includes confusion matrices, which cross-tabulate WSP integrated statuses vs. rapid statuses assigned by the algorithm, for all cases in the *learning data set*.

3.2 RECOMMENDED ALGORITHM: LEARNING TREE 3

We assessed candidate algorithms qualitatively and quantitatively according to the criteria outlined in Section 2.3.2. No algorithm excels on all six criteria. However, *Learning Tree 3* outperforms the other algorithms on the most criteria overall, due to its broad applicability, high accuracy, biologically conservative metric thresholds, and consistency with the common rationale used by experts in the WSP integrated status assessments.

Learning Tree 3 outperformed all the other algorithms on both number completed and number correct. The *Learning Tree 3* algorithm assigned a WSP rapid status to all 65 cases (100% completed) and has the highest overall number of correct completed assignment (54 cases, 83%; on the 3-status scale) (Table 7).

3.2.1 Criterion 1: Low Error Rate

Criterion1: Algorithms should have relatively low error rates when comparing WSP rapid statuses to integrated statuses, the latter which are assumed to be 'true' statuses.

Learning Tree 3 performed best when each algorithm is compared to the most appropriate status scale (5, 3, or 2 status levels), achieving 54 correct on the 3 status scale out of 65 cases (grey shaded cells in Table 7). *Learning Tree 3* also achieved the highest number of 'close' status assignments on all three status scales (within 1 step on the error scores in Table 4). The 2 status scale is most appropriate to evaluate performance of Simply *Red*. However, even on this scale, *Learning Tree 3* has the greatest number correct (Table 7). On the 5 status scale *Fancy Pants* has the greatest number correct (47), higher than the *Learning Tree 3* (44), though its completion rate is lower (54/65; 83%) than the *Learning Tree 3* (65/65; 100%) (Table 7).

Overall, *Learning Tree 3* assigns almost as many CUs correctly as *Fancy Pants* even on the 5-point scale that *Learning Tree 3* was not designed for (44 vs 47). *Fancy Pants* assigns some of the mixed status CUs correctly but fails to classify many of the other CUs with less data.

The *Learning Tree 3* algorithm had the greatest number correct for sockeye (Table 8) and Chinook (Table 9), whereas for coho the *Minimalist, Fancy Pants* and *Categorical Realist* algorithms (Table 10) had the most correct cases. Though even for coho, *Learning Tree 3* still had a high number of correct cases (4 out of 5).

Learning Tree performance improved from version 1 to version 3, in number completed and also number correct (Table 7). This supports our decision to combine information from the fitted trees with information from common rationale generated in the expert workshops. It also highlights the importance of continuous review and tweaking of the *Learning Tree* algorithm as additional case studies are completed.

3.2.2 Criterion 2: Precautionary

Criterion 2: Algorithm errors should be precautionary, meaning that estimated rapid statuses should err on the side of being poorer, indicating a higher risk of extirpation, when compared to 'true' integrated WSP statuses. For example, if a 'true' integrated WSP status is Amber, a status error should be more likely to be Red over Green.

Algorithms differed substantially in terms of the total number of over-predicted cases, but for this criterion the completion rate needs to be considered as well (Table 7). For example, Learning Tree 3 assigned better status than the expert consensus for 7 out of 65 completed cases (11%). Fancy Pants, on the other hand, only over-predicted 1 out of 54 completed cases (2%), which is a much better performance on this single criterion. However, Learning Tree 3 assigned more cases correctly (Criterion 1) than Fancy Pants (54 vs. 49 on the 3 status scale) and completed more cases (Criterion 3).

3.2.3 Criterion 3: Broadly Applicable

Criterion 3: Algorithms must be broadly applicable across CUs with different data types and metric availability.

The number of completed cases for the entire *learning data set* varied from 54 to 65 out of 65. The *Learning Tree* algorithms have the highest completion rate, each assigning statuses to 100% of the *learning data set* (Table 7).

Completion rate differed by species (Table 8 to 10) because data types and metric availability differ by species. Chinook CUs are much more limited in metric availability than coho and most sockeye CUs. Most (14 out of 15) Chinook CUs are missing both the *relative abundance* and *absolute abundance* metrics, while only four sockeye CUs are missing both of these metrics, and 11 sockeye CUs are missing the *relative abundance* metric only.

For Fraser sockeye, the three *Learning Tree* algorithms performed the best, with all 45 cases completed, and *Categorical Realist* performed the worst, with only 35 cases completed (Table 8). Most of the algorithms completed all 15 cases for Southern BC Chinook (Table 9), apart from *Fancy Pants* and *Simply Red* (11 cases completed each).

All algorithms completed all 5 cases for Interior Fraser coho (Table 10).

3.2.4 Criterion 4: Three Status Zones

Criterion 4: Algorithms that estimate WSP rapid status for three main status zones, Red, Amber, and Green are preferred.

Four of the seven algorithms assign status on the 3 status scale: *Minimalist*, and *Learning Tree 1-3.*

While *Fancy Pants* assigns status on the 5 status scale, the resulting status could be simplified to the 3 status scale (i.e. by converting any *Red/Amber* to *Red*, and any *Amber/Green* to *Amber*).

Categorical Realist and *Simply Red*, however, assign only two status zones and do not meet this criterion. This was intentional for *Simply Red*, which was specifically designed to assign either *Red* or *Not Red*. It was a result of the CART fitting for the *Categorical Realist*, which only has branches assigning either *Amber* or *Red*. For both algorithms this characteristic restricts the ability to capture the full range of status required by WSP rapid status applications.

3.2.5 Criterion 5: Status Thresholds Consistent With Published WSP Assessments

Criterion 5: Algorithms should reflect thresholds that emerged as those distinguishing statuses in WSP integrated status assessment. These tend to be equal to or more biologically conservative than WSP benchmarks for individual metrics from Holt et al. (2009).

The CART-fitted algorithms, and the *Simply Red* constructed algorithm, which was assembled from parts of the fitted algorithms, use fitted thresholds to distinguish tree branches (e.g. for the *long-term trend* metric). The *Learning Tree* algorithms combine these fitted thresholds with a review of the CU status narratives, to ensure that the thresholds are consistent with WSP intent.

The only algorithms that apply WSP benchmarks directly are the *Categorical Realist* and the *Learning Tree 2*, which use metric statuses instead of values to determine WSP rapid status. However, in *Learning Trees 1* and *3* metric thresholds derived from the *Minimalist* and *Fancy Pants* algorithms were adjusted to better align with the WSP benchmarks, and also to better fit the data, and align with the common rationale used by experts to assign status during the WSP integrated status assessments. As mentioned, metric thresholds in *Learning Tree 1* and *Learning Tree 3* are generally more biologically conservative than the WSP benchmarks, with the exception of one tree branch (Section 2.4.3).

3.2.6 Criterion 6: Rationale Consistent With Published WSP Assessments

Criterion 6: Algorithm decisions should adhere to the logic applied in the WSP integrated status assessments. This includes following common rationale applied in the detailed WSP status assessment processes, as documented in the CU status narratives reprinted in Appendix B, which includes extracts from Grant and Pestal 2012, Grant et al. 2020, Brown et al. (2014) and Parken et al. (2014).

Only the *Learning Tree* algorithms were developed with explicit consideration of the overall and CU-specific common rationales used by experts to assign WSP integrated statuses (Appendix B). Basically, the *Learning Tree* algorithms incorporate both status considerations revealed in the 65 completed assessments as well as the stated broader rationale in the workshop consensus. The fitted trees only draw on the quantitative metrics and status designations. The *Learning Tree* algorithms have also been carefully vetted by experts for logical flow.

3.3 CONFIDENCE RATINGS FOR LEARNING TREE 3 STATUS RESULTS

WSP rapid statuses assigned by *Learning Tree 3* were categorized as *High, Medium,* or *Low* for most of the 65 cases in the *learning data* set (Table 12). These confidence ratings were compared to the errors between the WSP rapid statuses and the WSP integrated statuses. Given the high overall success rate of this algorithm, there were only a few errors where WSP rapid statuses and integrated statuses did not align. Specifically, there were five cases where the algorithm assigned a better status than the WSP integrated status, and did so with *High* confidence (Table 12). These are the outcomes we would want to minimize, since they are less precautionary.

In all five cases, the discrepancy between WSP rapid and integrated statuses was small, and can be readily explained by the additional information considered in the expert deliberations, as documented in the status narratives for each CU within the WSP assessment reports (Appendix B). In all these cases, *Learning Tree 3* generates WSP rapid statuses that match the starting point for the experts' integrated status discussions, but the algorithm cannot capture the nuances of additional information used in the workshop consensus to downgrade the statuses by half a category (i.e. from *Amber* to *Red/Amber* or from *Green* to *Amber/Green*). On the 3-status scale, these half-category statuses were then simplified to the lower status (i.e. *Red/Amber* became *Red*, *Amber/Green* became *Amber*), and therefore these show up as a full category error in the comparison.

Details for the five cases are:

Francois-Fraser sockeye (2010): Experts in the WSP integrated status assessment process classified this CU as *Red/Amber*, which is converted to *Red* on the 3 status scale to calculate error. *Learning Tree 3* assigns a WSP rapid status of *Amber* due to the *relative abundance* metric, where the recent generational average is between lower and upper benchmarks. So in part, this discrepancy comes from converting the WSP integrated *Red/Amber* status to *Red* on the 3 status scale to calculate error. In the WSP integrated status assessments, experts used recent abundance and the *long-term trend* metric to move towards *Amber*, but then downgraded status based on additional information: low productivity (R/S) and sensitivity of the *relative abundance* metric (i.e. *Red* on some probability levels of the lower benchmark S_{gen} estimates for the time-varying productivity model).

Francois-Fraser sockeye (2015): Experts in the WSP integrated status assessment process classified this CU as *Amber/Green*. *Learning Tree 3* assigns a WSP rapid status of *Green* because the recent generational average was more than 10% above the upper *relative abundance* benchmark (i.e. RelAbd > 1.1*UBM). However, in the WSP integrated status assessments assessment the experts considered the full Bayesian posterior distributions for a suite of alternative benchmark estimates, as well as the retrospective pattern, both of which showed a mixture of *Red* and *Amber* metric statuses, and therefore the expert consensus was an *Amber/Green* designation. The algorithm uses a simplification of this, based on the median of a single benchmark estimate.

Shuswap-Late sockeye (2015): The expert workshop classified this CU as *Amber/Green*. *Learning Tree 3* assigns a WSP rapid status of *Green* because generational average was more than 10% above the upper *relative abundance* benchmark (i.e. RelAbd > 1.1*UBM). Experts in the status workshop used additional information to downgrade the status: (1) low abundance, (2) declining trends on off-cycle years, (3) *percent change* and *long-term trend* were given less weight in the expert workshop due to higher uncertainty in off-cycle estimates (last years in the time series). Note that the trend metrics are not used for this CU in the *Learning Tree 3* algorithm, because the branch it lands on focuses on *relative abundance*.

Takla-Trembleur-S sockeye (2015): The expert workshop classified this CU as *Red/Amber*. *Learning Tree 3* assigns a WSP rapid status of *Amber*. The experts considered a broader suite of information to downgrade status: different probability level of the *relative abundance* benchmark statuses rather than just the median, and cycle-line abundance, in addition to the generational average. In addition, the experts highlighted the steep recent decline (*percent change*) and declining productivity (R/S) to downgrade the status assessment.

North Thompson coho (2013): The expert workshop classified this CU as *Amber/Green*. *Learning Tree 3* assigns a WSP rapid status of *Green* because generational average was more than 10% above the upper *relative abundance* benchmark (i.e., RelAbd > 1.1*UBM). Experts in the status workshop used additional information with mixed signals to downgrade the status: *percent change* was increasing, *long-term trend* was in the *Amber* or *Red* zone in recent years of the retrospective, productivity (R/S) was below replacement in recent years, and marine survival was low but stable.

Discrepancies between WSP rapid statuses and WSP integrated statuses indicate potential factors to consider in future versions of the algorithm (e.g. uncertainty in benchmark, trends in off-cycle lines, productivity, marine survival). However, it will be a challenge to develop standardized metrics for these factors that are applicable across data sets.

3.4 PERFORMANCE IN THE RETROSPECTIVE (OUT-OF-SAMPLES) TEST

Status is intended to communicate the overall signal in the data. We tested algorithms retrospectively to evaluate the stability of statuses over time, as spawner abundances and resulting status metrics change, and across algorithms (Appendix G; Appendix H). The seven candidate algorithms were applied retrospectively to all CUs in the *retrospective data set* for years from 1995 to 2019, with applicable data. Figure 17 illustrates the resulting pattern in WSP rapid statuses for the *Learning Tree 3* algorithm for Fraser sockeye CUs.

Overall, this retrospective test generated 871 individual cases for which at least one metric was available for a CU, and a WSP rapid status could potentially be calculated. The total number of CUs that can be assessed increases over the course of the time series, as more data become available, and more metrics can be calculated (Figure 18 and Figure 19). This was consistent across algorithms despite large variation in the specific number of statuses completed by each.

Statuses generally changed over the time series, but were relatively stable, staying on one status zone for one or more years, before changing to another status zone. There are examples, like Widgeon_RT (river-type) sockeye where the WSP rapid status stayed *Red* over the entire time series, since this CU has a small geographic distribution, and therefore, small numbers that make it more vulnerable to extinction.

Patterns in the retrospective analysis show that the algorithms detect major population signals that are known to salmon experts. This illustrates how the algorithm results will be used within DFO's Salmon Scanner to examine summary patterns over time, and across species/areas. Specifically, the total number of *Red* CUs increases over time across all algorithms (Figure 18). This pattern is interrupted in the early 2010s due to the large returns observed for several Fraser sockeye CUs in 2010, which improves some CU statuses for several years (Figure 18).

The consistency among algorithms in identifying CUs with *Red* status is reassuring for future applications of the WSP rapid status approach. This shows that the overall pattern in WSP rapid statuses is not highly sensitive to the technical details of the algorithms. A key reason for this is that participating experts in the WSP status workshops generally reached consensus relatively easily on CUs with clear-cut indications of poor status, and those consistent considerations could then be extracted by the algorithms.

Appendix G and Appendix H document detailed results from the retrospective test. Notable observations are:

- Appendix G, Table 49, shows the competition rate across years by CUs where a WSP rapid status could be estimated. It specifically looks at how completion rate compares based on number of algorithms (from 0 to 7), and number of metrics (from 0 to 4).

Metrics include relative abundance, absolute abundance, long-term trend, and percent change. An obvious result of the retrospective test is when there were 4 metric values that could be estimated, then all seven algorithms could assign status (total number of cases: 509). Conversely, when no metric values could be estimated, then there were no algorithms that could assign status (total number of cases: 316). For a varying number of metric values that could be estimated, there were a varying number of algorithms that could estimate WSP rapid status.

- Appendix G, Table 50, also shows the completion rate across years by CU where a WSP rapid status could be estimated. However, it specifically looks at how completion rate compares across each of the seven candidate algorithms. There are pronounced differences in completion rate between algorithms. *Learning Tree 3* was the only algorithm that could assign status to all 841 cases with two or more of the four standard metrics, outperforming all other algorithms. In cases where there were four metrics, completion rate was 100% for all seven algorithms, though among all Pacific Salmon CUs, a small proportion will have all four metrics.
- The retrospective test identified CUs for which several of the candidate algorithms indicate a deteriorating status since the past WSP integrated status assessments (Appendix G.2). This includes two of the five coho CUs (Fraser Canyon coho, and North Thompson coho) and several sockeye CUS (5 changed from Amber to Green: Chilko-S-ES, Francois-Fraser-S, Pitt-ES, Shuswap-ES, Shuswap-L; 5 changed from Amber to Red: Chilliwack-ES, Kamloops-ES, Lillooet-Harrison-L, Nahatlatch-ES, North-Barriere-ES; and Harrison (River-type) changed from Green to Red.
- Overall, the retrospective pattern of WSP rapid statuses was generally consistent for a CU across algorithms and years (Appendix H). What we mean by patterns is while the exact status might not be identical, directional changes in status such as improvements or deterioration in status are similar. Actual WSP rapid status designations could vary between algorithms for some years and CUs.
- Across CUs patterns in WSP rapid statuses across time varied. Some CUs did not vary markedly over time, for example large and stable CUs like Pitt-ES, and were more consistently *Green* over their time series, and conversely small CUs like Taseko-ES or Widgeon river-type were consistently *Red.* Others showed varying statuses over time.

3.5 PERFORMANCE IN THE RELATIVE BENCHMARK METRIC SENSITIVITY TEST

The *relative abundance* metric sensitivity test compared how WSP rapid statuses changed when this metric was included or excluded from a candidate algorithm Table 11. The *relative abundance* metric was available for 37 of the 65 cases in the *learning data set*. Details on what CUs included this metric are in Appendix C. Only *Learning Tree* 1,2 and 3 algorithms sensitivity test results were considered further.

Two other algorithms can include the *relative abundance metrics: Categorical Realist and Simply Red*. The *Categorical Realist* and *Simply Red*, however, are relatively simple and cannot assign statuses for most cases, respectively 37/37 cases and 25/37 cases, once the *relative abundance* metric is excluded. Therefore, these are not discussed further below.

The *Minimalist* does not use the *relative abundance* metric, so it was excluded. The *Fancy Pants* algorithm ran into challenges with special cases that could not be addressed with the current code. However, given the observed completion rates and errors (summarized above), *Fancy Pants* is not the recommended algorithm, so we did not dedicate effort to further adjust the code. Therefore, we first present the results of this sensitivity test only for the three versions of the *Learning Tree*, then conclude with a comparison performance of *Learning Tree 3* without *relative abundance* metric to the much simpler *Minimalist* algorithm*.*

Completion rate was almost 100% for the three *Learning Tree* algorithms for this sensitivity analysis. There was only 1 unique case where statuses could not be assigned when the *relative abundance* metric was excluded (Table 11: *1 case in the Without Benchmark, Number Not Completed Row). This case is a special situation, because the initial WSP integrated assessment of Chilliwack-ES sockeye relied only on the *relative abundance* metric to assign status. At the time of this assessment, there was not sufficient data to calculate either *percent change* or *long-term trend* for this CU (Grant & Pestal 2012). The *relative abundance* metric for Chilliwack-ES required only the most recent generation of escapement data (last four years), because benchmarks were calculated using the rearing capacity of Chilliwack Lake (20% and 40% of spawners at maximum juvenile production (S_{max})) (Grant et al. 2011). When the *relative abundance* metric was excluded, there was no other information available to assess status (Table 11: *1 case for *Learning Trees*).

As the *Learning Tree* evolved from 1 to 3, the algorithm became more consistent in the statuses assigned with the *relative abundance* metric versus without. Statuses changed when the *relative abundance* metric was removed for the following number of cases: *Learning Tree 1*: 17/37 cases; *Learning Tree 2*: 13/37 cases; *Learning Tree 3*: 9/37 cases (Table 11). *Learning Tree 3* is more precautionary when the *relative abundance* metric is not available, compared to when it is available, and compared to *Learning Trees 1* and *2*. In most cases (7/9), the WSP rapid status assigned with less information (i.e. without the *relative abundance* metric) was poorer than with it included (Table 11: Number Worse by 1 status zone: 5; Number Worse by 2 status zones: 2; out of the total Number Changed: 9).

The opposite occurred with *Learning Tree 1*, where in most cases (13/17), the WSP rapid status assigned without the *relative abundance* metric was better than with it included (Table 11: Number Better by 1 status zone: 13; Number Better by 2 status zones: 1; out of the total Number Changed: 17). This is notable, because the metrics and thresholds used by *Learning Tree 1* and *3* are very similar (see Section 2.4.3 for differences between *Learning Tree 1 and 3*).

When constructing *Learning Tree 3,* we added a branch (at node 32, see Appendix E.7), so that one portion of the tree resembles the *Minimalist* algorithm. The *Minimalist* algorithm performed very well in terms of Number Correct and Number Close on the 3 status scale (tied for #2 rank across algorithms), despite being very simple and relying only on *trends in abundance* metrics. Cases that end up on this branch of the *Learning Tree 3* are those CUs that have no *relative abundance* metric, and either have no absolute abundance data, or have absolute abundance estimates that are larger than 10,000.

In the relative benchmark sensitivity test, 21 CUs followed the branch of the *Learning Tree 3* that resembles the *Minimalist.* For 19 of those 21, algorithm results matched those produced by the *Minimalist*. For the remaining two cases, *Learning Tree 3* gives a different result, assigning *Red* statuses where the *Minimalist* assigns a WSP rapid status of *Amber*. For these 2 CUs the *Learning Tree 3* (without the *relative abundance* benchmark) actually performs worse (0 correct statuses) than the *Minimalist* algorithm (1 correct status). However, this is because the *Learning Tree 3* algorithm is intentionally more cautious for consistency with

COSEWIC criterion A (as described in Section 2.4.3).

Specifically, the *Learning Tree 3* assigns a *Red* status for CUs that show a 70% decline on the *percent change* metric, while the *Minimalist* algorithm requires an 80% decline on this metric to assign *Red* status.

4 DISCUSSION

4.1 WSP RAPID STATUSES

4.1.1 Selected Algorithm: Learning Tree 3

Learning Tree 3 (Figure 20) was selected to be used as the WSP rapid status algorithm going forward. This algorithm performed best across the seven candidate algorithms (See Section 4.1.3.1 below). The *Learning Tree 3* assigns a *Red, Amber* or *Green* WSP rapid status, with a *High, Medium,* or *Low* confidence rating for CUs with applicable data.

The *Learning Tree 3* algorithm consists of a set of decision rules that approximate the decision-making process that experts used to assign CU statuses in past WSP integrated status assessment processes (Figure 20). It assigns a WSP rapid status depending on answers to a series of Yes/No questions. The algorithm sequence is as follows:

1. The first question is whether or not a CU has a current *absolute abundance* value, and if so, whether or not this value falls below the lower threshold of 1,500 (which adds a buffer to COSEWIC's Criterion D1 for small population size of 1,000). If the answer to this question is Yes, then the CU is assigned *Red* (node 3), with *High* confidence.
2. If the answer to the first question is No, then the second question is whether or not the CU has a current *absolute abundance* value, and if so, whether or not the current abundance is below the upper threshold of 10,000, which is COSEWIC's Criterion C upper benchmark. This second question splits the decision nodes into two Pathways: **Pathway 1 (No to this question)** and **Pathway 2 (Yes to this question).**
    - **Pathway 1:** is where a CU either does not have a current absolute abundance value, or has these data, and it falls above the upper threshold for this metric. This pathway is split with the question: can this CU be assessed with a *relative abundance* metric. If the answer is Yes, a *Red* (nodes 19), *Amber* (nodes 37) or *Green* (node 36) WSP rapid status is assigned, with *High* confidence, depending on where the current abundance value falls relative to this metric's lower and upper thresholds. If the answer is NO, then comparisons are made between the CU's current abundances and percent change to thresholds for these metrics, which assign a *Red* with *Medium* confidence, or *Green* or *Amber* with *Low* confidence status.
    - **Pathway 2**: is where a CU has absolute abundance data, and these abundances fall between the lower and upper thresholds. In this pathway, absolute abundances restrict WSP rapid statuses to only *Amber* or *Red*. This pathway is split with the question: can this CU be assessed with a *relative abundance* metric. If the answer is Yes, an *Amber* (node 22) with *Medium*

confidence, or *Red* (node 23) with *High* confidence, is assigned, depending on whether the CU's current abundance value falls above the relative abundance metric lower threshold or below. If the CU cannot be assessed with a *relative abundance* metric, then it is compared to the lower threshold of the Long-Term trend metric and assigned *Amber* (node 20) with *Medium* confidence if above, or *Red* (node 21) with *Medium* confidence if below.

4.1.2 Fitted CART Algorithms: Starting Point For Algorithm Development

The fitted algorithms using CART models were a useful starting point in the algorithm development process, which concluded with *Learning Tree 3*. They helped us explore the range of algorithms that could be derived from the *learning data set*, from those that produce simple status results (*Categorical Realist*: *Red* and *Not Red*) to the full portfolio of status results (*Fancy Pants*: *Red*, *Red/Amber*, *Amber*, *Amber/Green* and *Green*), with one intermediate between these bookends (*Minimalist*: *Red*, *Amber* and *Green*) (Table 6).

However, the fitted algorithms had several limitations. First, the small number of *learning data set* cases, which included only previously completed WSP integrated status assessments, limited the number of patterns the CART decision trees could extract.

In addition, the *learning data set* is not balanced across species, with 45 sockeye cases, 15 Chinook cases, and only 5 coho cases. This means that CART-fitted algorithms may be overfit to sockeye CUs, for which we have more cases, and therefore, perform less well for Chinook and coho CUs, for which we have fewer cases. This is particularly important because the sockeye CUs differ from both Chinook and coho in terms of data quality, length of CU time series, and the weighting of metrics in the WSP status assessments.

More metrics can typically be calculated for Fraser sockeye CUs than for Chinook CUs, and metrics for sockeye were generally considered well estimated in the WSP integrated status assessments, due to the availability of long, high quality data sets for many Fraser sockeye CUs. This translated into greater weight being placed on the *relative abundance* metric in the sockeye process than in the coho or Chinook assessments, where this metric was often down-weighted (coho) or not available (Chinook). Algorithms that are well fit to the *learning data set* therefore tend to place greater weight on the *relative abundance* metric than is appropriate for coho and Chinook. Finally, sockeye CUs also have very different abundance patterns than the coho CUs, with many having peaked in abundance in the late 1990s and early 2000s followed by declines. This influenced the weight attributed to the *percent change* metric in the WSP integrated status assessments for many sockeye CUs, which was more heavily relied upon for coho and Chinook.

The differences in metric weighting across WSP integrated status assessments for different species, historical trends, and data types are highly nuanced, and are not easily captured through the CART fitting process, particularly given the distribution of cases in the data set.

The fitted algorithms, therefore, did not represent an end point for algorithm development and selection. Instead, they provided a useful starting point for exploring and comparing the performance of a range of classification trees to support the development of constructed trees.

4.1.3 Constructed Algorithms: Concluding With Learning Tree 3

4.1.3.1 Learning Tree development, performance, and broad applicability to BC and Yukon CUs

The *Learning Tree* algorithms were the final evolutionary step in developing the constructed algorithms. They were derived by combining components of CART-fitted trees with common rationale applied in the expert-driven WSP status assessment processes. This set of algorithms was developed iteratively as WSP status and CU experts incorporated and refined decision rules and metric thresholds and compared changes to the algorithm's performance. The *Learning Tree 3* (Figure 20) is the culmination of this iterative revision process. Already, *Learning Tree 1* evolved into *Learning Tree 3*, improving in accuracy and over-prediction error as additional considerations were added. No further improvements in performance were found once the *Learning Tree 3* algorithm version was reached.

The *Learning Tree 3* performed best overall. It is applicable to the largest proportion of CUs in the *learning data set* (100% of cases), it has the highest accuracy (83% correct overall on the 3-status scale, 84% Fraser sockeye, 80% for SBC Chinook and Interior Fraser coho), and it adheres to the decision-making processes that occurred in the WSP integrated status assessments, including applying biologically conservative metric thresholds. Given that the *Learning Tree 3* algorithm has the highest completion rate for the *learning data set*, it should also be the algorithm that is most widely applicable to other species and areas.

Learning Tree 3 was designed to account for differences in type of data (i.e., relative index versus absolute abundance), and suite of metrics available. *Learning Tree 3* provides branch options that are conditional on metric availability. This flexibility ensures its applicability across BC and Yukon CUs, where many CUs have abundance data that are indices only, and therefore, only *trends in abundance* metrics will be applicable.

The name *Learning Tree* was deliberately chosen to indicate that this algorithm can continue to improve over time. As new metrics and CUs are considered, additional WSP integrated status assessments are recommended. This will enable any necessary adjustments to the *Learning Tree 3* algorithm, by expanding the *learning data set* to evaluate the performance of new *Learning Tree* algorithm adjustments, relative to previous versions.

The only other constructed algorithm *Simply Red*, is limited in its applicability across the range of data types and metric availability within the *learning data set*, due to its reliance on *relative abundance* and/or *absolute abundance* metrics to assign status. This algorithm also had among the lowest completion rate overall due to this limitation.

4.1.3.2 Learning Tree Error

The *Learning Tree 3* algorithm did a good job at replicating the expert decision-making processes of past WSP integrated status assessments (Grant & Pestal 2012; DFO 2015; DFO 2016; Grant et al. 2020). There are, however, nuances in the expert decision-making processes that were unique for particular CUs and could not be generalized within the *Learning Tree 3* algorithm. This reflects the trade-off between having an algorithm that can rapidly and annually assess status for all Pacific Salmon CUs with applicable data, versus more detailed processes that only can assess a small percentage of CUs for single years but can address these nuances through expert input.

Nuances that could not be generalized in the *Learning Tree* algorithms include:

- recent trends in productivity;
- uncertainty in the *relative abundance* metric benchmarks, which placed the metric in additional status zones when presented across a range of probability levels;
- retrospective values of the key metrics;
- specific cycle-line trends in the cases of some cyclic sockeye CUs; and,
- inclusion of the *relative abundance* metric in decision-making for some Chinook CUs, where we have chosen to exclude this metric due to its inappropriate application given the data issues of the CU.

4.1.3.3 Confidence in WSP Rapid Status

We incorporated a confidence rating for *Learning Tree* 3 statuses using three confidence categories: *Low, Medium* or *High* (Figure 20). This confidence rating largely addresses the fact that even expert consensus on a status designation, developed in a workshop setting, will be associated with higher or lower confidence, depending on the type of CU information available to assess status. Lower confidence in expert driven processes led to more divergence among experts in regards to their initial status designations (Figure 3 and Figure 4). For this reason, you cannot identify confidence using errors in the *learning data set* for the *Learning Tree* algorithm. WSP rapid status errors were associated largely with the additional information that experts used to assess WSP integrated status for particular cases, as opposed to confidence in the metrics and data available to assess status.

High confidence statuses are generally those that are assigned using *relative abundance* and/or *absolute abundance* metrics. *Low* confidence statuses rely exclusively on *long-term trend*, and may also include the *percent change* metric, and provide statuses of *Green* or *Amber*. *Medium* confidence statuses include nuances between these two categories.

4.2 CHANGES IN STATUS SINCE THE LAST WSP INTEGRATED STATUS ASSESSMENTS

Four integrated status assessments under the WSP have been completed; two for Fraser sockeye, and one for Interior Fraser coho, and one for Southern BC Chinook. These assessments covered 47 CUs from three species of salmon. *Learning Tree 3*, the recommended algorithm, indicates changes in status for many of these CUs since their last formal integrated assessment, using available data up to 2018 or 2019, depending on the CU. The WSP rapid statuses show a deterioration since the last formal assessment for 11 of the 23 Fraser sockeye CUs, and for 4 of the 15 Southern BC Chinook CUs with enough data from wild sites to complete an assessment (Appendix G; Appendix H). The number of CUs with a *Red* status increased from 1995 to 2019 (Figure 18). Conversely, the percentage of CUs assigned a *Green* or *Amber* declined over time (Figure 19). This confirms the urgent need for up-to-date status assessments and demonstrates the usefulness of the recommended algorithm.

We are documenting the details of these status changes in a companion report, which applies the recommended *Learning Tree 3* algorithm to the latest available data sets for Fraser sockeye, Interior Fraser coho, Southern BC Chinook, Fraser pink, Fraser chum, Skeena sockeye, and Nass sockeye CUs (Pestal et al. 2023).

4.3 LAYERS OF PRECAUTION

We chose to be precautionary at multiple stages of the WSP rapid status algorithm process to align this approach with the WSP integrated status assessment approach, which provides 'true' CU statuses. Precautionary actions taken were:

1. To evaluate algorithm performance, we downgraded mixed WSP integrated statuses to the poorer of the two statuses (*Red/Amber* became *Red*, *Amber/Green* became *Amber*).
2. In the evaluation of alternative algorithms, we looked at the direction of errors, and considered underestimates of status (e.g. assigning *Amber* status to a *Green* CU) less of a concern than overestimates of status (i.e. Criterion 2, Section 2.3).
3. In the *Learning Tree 3* algorithm:
    - o we included a buffer of 500 above the COSEWIC absolute abundance Criterion D1 threshold of 1,000 for small population size; the threshold for this metric is set at 1,500. This was to account for how this metric was treated by experts in the workshops, where CU statuses were downgraded if one year in a generation fell below 1,000, if the estimates were considered uncertain, or if the generational average was close to the 1,000 threshold.
    - o similar to the buffer on the *absolute abundance* metric lower benchmark (previous bullet), we added a 10% buffer to the upper threshold of the *relative abundance* metric.

This level of precaution in the WSP rapid status assessment approach is consistent with IUCN and COSEWIC status evaluation approaches (Mace et al. 2008). Both IUCN and COSEWIC status assessments are precautionary, which can result in some over listing: i.e. including a wildlife species in a threat category such as *Endangered*, *Threatened*, or *Special Concern,* when it is close to status thresholds that delineate these statuses categories, from *Not at Risk.* These 'at risk' designations flag species for urgent closer inspection and diagnosis, to determine if conservation actions are required. The alternative, less risk averse approaches, increases the risk of misclassifying a species in a *Not at Risk* category in error, when it is, in fact, facing an increased risk of extinction. This would result in a critical missed opportunity to initiate conservation actions in time to prevent the species' extinction.

The WSP rapid statuses are similarly designed to flag potential concerns that are meant to be further explored in subsequent evaluation processes (Section 4.4). We therefore consider it appropriate that WSP rapid statuses err on the side of caution, raising a flag in borderline cases. If a CU is assigned a WSP rapid status of *Red* or *Amber*, this is equivalent to the 'check engine' light coming on in a car. There are many potential reasons for the warning, and how you respond to the warning will depend on the details of the situation. But the first step after being flagged by the algorithm is to have a closer look at these cases.

The level of precaution in the WSP rapid status assessment approach is also consistent with the approach taken by experts in the completed WSP integrated status assessments (Appendix B; Grant & Pestal 2012; DFO 2015; DFO 2016; Grant et al. 2020).

Examples of where experts in the WSP integrated status assessment processes included precautionary approaches are provided below:

- In the WSP integrated status assessment processes, the *relative abundance* metric drove status designations where it was available. In evaluating this metric, experts considered the consistency in status across all probability levels (10% to 90%) of the estimated benchmarks to determine status. If statuses were mixed across probability levels, status was down-weighed towards the lower status level, or a mixed status was assigned (e.g. *Red/Amber* or *Amber/Green*) (Appendix B). The WSP rapid status approach compares the current generational average (or 'dominant' cycle, in the case of Fraser sockeye cyclic CUs), to the median (50% probability level) estimates of the *relative abundance* benchmarks, instead of presenting the full probability distribution of the benchmarks. Since this metric is so heavily relied upon in status designations, using only the median benchmarks in the WSP rapid status algorithm has the potential to assign overly optimistic statuses in comparison to the WSP integrated status approaches. The three decisions listed above were therefore made to remain consistent with the degree of caution applied in the expert-driven processes.

- When considering absolute abundance in WSP integrated status workshops, experts considered uncertainty in the data, and also compared each of the past four to twelve years to the COSEWIC criterion D1 (small population size) threshold of 1,000. In contrast, the algorithm compares the last generation average abundance to the COSEWIC threshold. To make this algorithm threshold more consistent with the precautionary approach used by experts in the WSP integrated status approach, a buffer was added. The buffer accounts for data uncertainty, and some of the masking of individual low abundance years (falling below the 1,000 COSEWIC threshold) that might occur, when averaged together with larger abundance years in the most recent generation. The buffer of 500 increases the COSEWIC metric threshold to 1,500 in the algorithm.

Note that biological thresholds for WSP rapid status are currently stationary. They do not consider deteriorating salmon productivity observed for many salmon CUs (Dorner et al. 2008, 2018; Grant et al. 2019; MacDonald et al. 2023). As the climate continues to change and habitats continue to deteriorate due to human activities, larger salmon population size thresholds may be required to ensure a CU's persistence under these conditions (McElhany et al. 2000). See next steps section below on consideration of time-varying productivity in the WSP rapid status approach.

4.4 FUTURE CONSIDERATIONS FOR THE WSP RAPID STATUS ALGORITHM (*LEARNING TREE* 3 ALGORITHM)

4.4.1 Summary

This final section on future considerations for the *Learning Tree Algorithm* is organized under the following:

1. **The second core principle of WSP rapid status assessment is the vetting of data and evaluation of statuses by CU experts that manage the data for specific groups of salmon CUs.** To address this principle we present two key elements:
    - Identify key data processing steps;
    - Develop a data management strategy for WSP rapid status assessments.

2. **The third core principle of the WSP rapid status algorithm is continual learning and refinement.** Refinements to the *Learning Tree 3* algorithm and how it is used can include the following (details in subsequent sections below):

    - Algorithm revisions as required. This includes changes to the algorithm and/or adding new metrics;
    - Adding or updating *relative-abundance* benchmarks for CUs; including incorporating time-varying productivity into the metrics;
    - Explore revisions to data sets with hatchery influence using the Proportionate Natural Influence (PNI) in salmon CU statuses;

3. **WSP rapid statuses and DFO's new Salmon Scanner Applications**
    - Improving End-User Access;
    - DFO's Scanner (Data-Visualization tool);
    - Applications of WSP rapid statuses and DFO's Salmon Scanner.

4.4.2 <u>The second core principle of WSP rapid status assessment is the vetting of data by CU experts that manage the data for specific groups of salmon CUs.</u>

4.4.2.1 <u>Key Data Processing Steps</u>

Data is selected and treated based on the expertise of DFO stock assessment staff, who work in collaboration with their DFO teams, and with Indigenous groups, consultants, NGO's, etc. Through the vetting process, we eliminate any data sets or metrics that would not produce reliable status results. This also ensures that data selection and treatment is standardized. This vetting step is required to label statuses 'WSP rapid statuses'.

We intend to apply the *Learning Tree 3* algorithm to additional data sets from BC and Yukon watersheds as these become available for use. Work is currently being initiated to apply the *Learning Tree 3* algorithm to CUs that do not have completed integrated status assessments. This includes Fraser pink and chum salmon, and Skeena and Nass sockeye salmon (Pestal et al. 2023). We will work with CU stock assessment experts to assign and evaluate statuses for these additional CUs and, if required, refine the *Learning Tree 3* algorithm by moving through the following steps:

1. Review CU-level data and calculate status metrics.
2. Review a range of input specifications like the start year of the time series, the generation length of the CU, metric applicability (e.g., "is the *relative abundance* metric meaningful for this CU, given the type of data and available SR estimates?") etc., as identified by CU experts.
3. Apply the *Learning Tree 3* algorithm with a range of input specifications, as recommended by CU experts, and review preliminary WSP rapid statuses.
4. Repeat steps 1-3 until there is consensus among the stock assessment experts that the WSP rapid statuses are reasonable.

It is important to note that the WSP rapid status algorithm approximates more detailed processes. The key with the algorithm is that it can be used to make relative comparisons between years within a CU, or across CUs by year given its standardized approach.

4.4.2.2 Develop a Data Management Strategy for WSP rapid status assessments

A key step in expanding WSP rapid status assessments is to develop a coordinated approach in DFO to manage the applicable salmon data. Due to DFO's Pacific Salmon Strategy, there is currently both increased resourcing, and an opportunity to put these pieces in place to ensure that all applicable Pacific Salmon CU data are available annually to assess WSP rapid statuses. Note, we assume DFO Area Stock Assessment leads integrate expertise from Indigenous groups, NGO's, consultants and others, in the management of CU stock assessment data. We recommend the following roles and responsibilities for Data Management consideration:

DFO Pacific Salmon Strategy Initiative (PSSI) Data Policy and Analytics Team

- **Creates and maintains central database (DB):** to warehouse annual composite data for WSP rapid status assessments, and annual CU WSP rapid status assessments and available WSP integrated status assessments; these data would be accessible to DFO staff and external groups: Indigenous groups, COSEWIC, IUCN, PSF, etc..

DFO Science: Data Management Unit (DMU)

- **Establish governance:** ensure annual CU composite data for WSP rapid status assessments, and WSP rapid statuses, are provided by DFO Stock Assessment leads.
- **Automate data treatment steps where possible:** this includes development of appropriate computer code packages and input specification files; in collaboration with PSSI Data Policy and Analytics Team and DFO Stock Assessment leads.
- **Ensure standardization in approaches across CUs and years:** work directly with Stock Assessment leads, and with support from State of Salmon Program (SOS) leads for new CUs.

Area and Core DFO Stock Assessment

- **Set up data treatment and specification files***:* for the WSP rapid status application for new CU data sets (following data steps in previous data section); in collaboration with SOS leads and DMU;
- **Provide annual selected and treated data***:* for WSP rapid status application to DMU.
- **Support the automation of data treatment steps where possible:** this includes development of appropriate computer code packages and input files; in collaboration with DMU and PSSI Data Policy and Analytics Team.
- **Support standardization processes across groups of CUs:** led by DMU.

DFO Science: State of the Salmon (SOS) Program (Authors of current paper)

- **Work with DFO Stock Assessment leads:** to determine data needs and metric specifications for WSP rapid statuses for new CU data sets being added (following data steps in previous data section); to ensure standardization across CUs and years.
- **Provide annual time series of WSP rapid statuses to DMU DB:** Pull data from data base and update rapid statuses.

4.4.3 The third core principle is continual learning and refinement of the WSP rapid status algorithm (Learning Tree 3)

4.4.3.1 Algorithm revisions as required

Revising the WSP rapid status algorithm can include directly altering the decision tree, or adding new metrics. Such revisions or improvements may be identified as new CU data sets are assessed. In such cases, we recommend that experts perform additional WSP integrated status assessments, to expand the *learning data set*. WSP integrated status assessment processes should include DFO and Indigenous groups and other experts, similar to past processes. With the existing or updated *learning data set,* performance of the WSP rapid status algorithm should be re-evaluated and compared between the existing algorithm and new algorithm revisions proposed. This would ensure that the algorithm's performance improves overall, when compared to the 'true' statuses, versus hyper-tuning the algorithm to particular CU cases.

As the number of CUs assessed through WSP rapid status assessments expands, there may be additional metrics that could be added to the *Learning Tree 3* algorithm. New metric considerations, however, should align with the WSP emphasis on 'standardized monitoring of [Pacific] salmon status' (DFO 2005; Holt et al. 2009). We also recommend continuing to emphasize standardized metrics and additional information that focuses on abundance and trends in abundance at this time (Appendix A; Holt et al. 2009; Holt 2009; Grant et al. 2011; Grant & Pestal 2012; DFO 2015; DFO 2016; Grant et al. 2020). These status metrics are based on conservation biology theory, particularly with emphasis on two paradigms: small population size and declining population (Caughley 1994; Mace et al. 2008).

A *distribution* metric currently is not included in the WSP rapid status algorithm. *Distribution* metrics were included in a WSP status toolkit (Holt et al. 2009), and CU distribution trends were provided in the Southern BC Chinook and Interior Fraser coho integrated status assessment processes. However, distribution information did not influence WSP integrated statuses (Appendix B; DFO 2015 & 2016). Further, no benchmarks have been resolved for *distribution* metrics through expert processes or research.

Distribution metrics might be particularly important to broadly distributed CUs, like those of chum and pink salmon. Considerable information on spawning distribution exists among salmon experts within DFO and among Indigenous communities and other groups. If work is done to develop benchmarks and explore their use by experts in WSP integrated status assessment processes, distribution metrics could be added to subsequent iterations of the *Learning Tree 3* algorithm. However, another important consideration is how broadly available these data will be across CUs, and how readily they can be updated annually.

Distribution information might be more relevant for subsequent steps involving the use of rapid statuses, rather than in the evaluation of status itself. For example, available information on changes within a CU's spawning or juvenile rearing distribution should be captured when developing recovery or rebuilding plans.

4.4.3.2 Adding or updating relative-abundance benchmarks for CUs, including incorporating time varying productivity into benchmarks.

Relative-abundance metric benchmarks should be added and updated for CUs where possible. These benchmarks are added by CU experts, based on their knowledge of the applicability of the data to this metric. Although WSP rapid statuses can be developed without *relative-abundance* metrics, the confidence in WSP rapid statuses increases when these metrics are available.

Broad declines in Canadian salmon abundances and productivity suggest that time-varying productivity should be considered in the *relative abundance* metric benchmarks. This is recommended for CUs where persistent changes in abundances and productivity have occurred. Time-varying productivity benchmarks, estimated from stock-recruitment models, were used in the first WSP integrated status assessment process for Fraser sockeye (Grant et al. 2011; Grant & Pestal 2012). However, these were not included in the subsequent WSP integrated status assessment since statuses of these CUs had returned to average, relative to the previous five years of poor productivity (Grant et al. 2020). Therefore, the more recent WSP integrated status assessment for Fraser sockeye CUs relied on models that considered average productivity for each CU (Grant et al. 2021). Since this last assessment, however, productivity declines have resumed. Further, since climate change is expected to continue to significantly change the quality of ecosystems and habitats, persistent CU productivity and distribution changes are expected (Bush and Lemmen 2019; Cheung and Frölicher 2020; IPCC 2021).

Incorporating time-varying productivity into relative abundance benchmarks is challenging when CU productivity has not stabilized (Peterman et al. 2003; Dorner et al. 2008, 2018; Malick et al. 2017), and when large productivity shifts continue to occur between years (Grant et al. 2021). Questions to consider include: how often to adjust benchmarks to account for time-varying CU productivity; how to interpret status over time if benchmarks are adjusted frequently, or are not adjusted despite productivity changes; and how to ensure consistency in applying time-varying productivity considerations to benchmarks in the WSP rapid status algorithm. Work is on-going in DFO to investigate these types of questions and develop guidelines in regard to developing and applying time-varying productivity to status and other applications such as forecasts (C.A Holt, DFO, pers. comm.).

4.4.3.3 Explore revisions to data sets with hatchery influence using the Proportionate Natural Influence (PNI) in salmon CU statuses

Hatcheries are expected to play an increasing role in the conservation of salmon CUs. Hatchery enhancement programs are being expanded for this purpose through DFO’s Pacific Salmon Strategy Initiative (PSSI). Although all WSP integrated status assessments to date have attempted to exclude hatchery populations (Grant et al. 2011; Brown et al. 2019), this may be increasingly challenging to do going forward given the larger role hatcheries will play in salmon conservation.

Recent work explores Proportionate Natural Influence (PNI) in hatchery influenced salmon (Withler et al. 2018). The PNI is a metric used to assess the genetic risks of hatchery production on natural populations as an index of gene flow. Guidance provided in a recent publication is being considered for adjusting which salmon populations should be included for a CU status assessment, depending on the level of PNI (see Table 3, in Withler et al. 2018).

4.4.4 Applications for WSP rapid statuses and DFO's new Salmon Scanner

4.4.4.1 DFO's Salmon Scanner: Improving End-User Access

WSP rapid statuses generated by the WSP rapid status algorithm have been incorporated into DFO's Salmon Scanner. This is an interactive data visualization tool for Pacific salmon. It is specifically designed for experts to support scientific discovery and help them contribute science to decision-making processes. Experts are those with expertise on Pacific salmon including stock assessment biologists, Indigenous groups, research scientists, habitat, harvest, and hatchery management biologists etc.

DFO's Salmon Scanner centralizes and makes WSP rapid statuses and key salmon data readily available to experts, including escapement, recruitment, life-history, and spawner distribution. DFO's Salmon Scanner enables technical experts to explore trends across CUs on different spatial, temporal, biological, and management-based scales. A key feature of DFO's Scanner is that it only includes quality-controlled data sets, prepared and vetted by CU experts for the purpose of WSP status assessments (see preceding data sections).

DFO's Salmon Scanner design process began over three years ago with structured questionnaires with 25+ experts across DFO Science and other management Sectors, academia, Indigenous groups. This was implemented to determine what these experts needed to do for their salmon-related work, but at the time were unable to. We then summarized the key tasks and priorities identified in the answered questionnaires to focus the design features of the Scanner. We used this information to create a basic Scanner version using Tableau, which is a visual analytics platform that facilitates prototyping. Initial design features were explored in this platform with experts by iteratively testing and making refinements. Design work was led by Dr. M. Barrus, an expert in software design.

After finalizing DFO's Scanner's design through this process, we re-developed in R-Shiny, and continued iterative testing and making refinements. R-Shiny provides considerably more flexibility to implement design features we identified through the expert-testing process. R-Shiny is a widely used freeware with many applications in fisheries science and decision support. We used DFO's Salmon Scanner in a 40+ person three day workshop to test various uses of this application. Like any software where regular updates and new releases occur, the Scanner will continue to evolve and change, as feedback is gained from on-going use.

In the coming months the current version of DFO's Salmon Scanner will be made available to DFO and external salmon technical experts, through individual and group sessions. It has been designed as a code package to be run on R, but can also be used in a browser format. To provide the source code for analysts, we have developed the *rapid_status()* function, which generates statuses using all the candidate algorithms, including the recommended *Learning Tree 3*. This code runs fast in R, and has been refined to handle many different types of special cases that may come up. The *rapid_status* function can be shared in several alternative formats: add it to the WSPMetrics R package available on Github; create a stand-alone RapidStatus package on Github; or offer a downloadable script. However, the analysis has a substantial learning curve in terms of setting up the data and interpreting the output, and may require an interactive application (e.g. in shiny) that guides users through each step and assists with interpretation. Setting up a basic app prototype could be fairly quick, but substantial design work and end-user testing would be needed to move from an app that works to an app that actually gets used, and used properly.

4.4.4.2 The Salmon Scanner Design Features

DFO's Salmon Scanner is divided into seven interactive tabs:

- **Filter Data Tab** (Figure 21): enables filtering data sets of interest by species, conservation unit, stock management unit, data availability, freshwater adaptive zone, life-history (such as stream-, river-, ocean-, lake-type), and average generation length. With this tab the expert can either choose to use the entire data set in subsequent tabs by not selecting anything, or narrow down specifically what they want to focus on by making selections.
- **View and Highlight on Map Tab** (Figure 21): this is the central component of DFO's Salmon Scanner. It enables experts to view the salmon watersheds throughout BC and the Yukon, through connected stream systems. CUs are mapped and can be colour-coded relative to WSP rapid statuses for user-specified year of interest, but can also be colour-coded relative to the stock management units, life-history traits, and more. The map can be interacted with, zoomed in and out, displayed in satellite mode, and other functionality end-users are familiar with from mapping applications. When CUs are selected, detailed time series and status information is presented in a lower panel under the map.
- **Time Series Plots Tab** (Figure 22): displays interactive figures that correspond to the filtered and selected CUs from the previous two tabs. This tab can produce publication or presentation figures that can be adjusted based on user specifications on colour, font size, labels, and more.
- **Compare CUs Tab:** only available in expert-user mode, this tab uses a parallel coordinates plot to enable the user to select CUs based on more detailed specifications by the user.
- **Table and Download Tab** (Figure 22): provides tabular access to directly interact with the data (e.g. alternative sorting) and download it for other applications.
- **Summary Reports Tab:** summarize the information the user has filtered and selected across freshwater adaptive zones, stock management units, etc.
- **Markdown Reports Tab:** provides pdf reports that summarize status trends for selected CUs over time; and provides one page CU summaries with plots of the time series, metrics and WSP rapid statuses.

4.4.4.3 Applications of WSP rapid statuses and DFO's Salmon Scanner

DFO's Salmon Scanner is designed for technical experts working on salmon. DFO's Salmon Scanner can be used by individual experts to support their work, by researchers to explore and develop hypotheses, by salmon management sectors to plan and evaluate outcomes of management actions, and by Indigenous groups, and others to support decision-making processes and other requirements.

In DFO's Salmon Scanner, WSP rapid statuses are a diagnostic tool to highlight potential issues and monitor trends across all of BC and Yukon Pacific salmon CUs, with sufficient data. This is particularly important as environmental conditions are broadly deteriorating due to climate change, and other human activities (Grant et al. 2019; IPCC 2022b).

DFO Science's State of the Salmon Program will use DFO's Salmon Scanner in expert processes to develop regular state of the salmon reports on salmon responses to changing conditions. Individual experts can also use DFO's Salmon Scanner to compare and contrast salmon CU statuses over time to determine how statuses are changing, and how CUs of interest compare to other CUs in BC and the Yukon. Scientists can explore broad hypotheses about trends in salmon statuses over time and space.

Hatchery, harvest and habitat experts can explore annual WSP rapid statuses and salmon abundance trends from DFO's Salmon Scanner to determine whether or not CU statuses are responding to management actions. This can support adaptive management to prioritize, improve or adjust, among their current practices.

DFO's Salmon Scanner can also be used to support rapid responses to emergencies like landslides, contaminant spills, forest fires, flooding, etc. For example, using the spatially-connected river systems mapped in DFO's Salmon Scanner, all CUs upstream of a particular mainstem location can be quickly selected and reviewed. Prior to DFO's Salmon Scanner there was no way to quickly access information on salmon that might be affected by an incident, such as the Big Bar landslide in the Fraser River that initially blocked salmon passage past the site (Government of B.C. et al. 2019). The Scanner fills this gap and also provides a efficient way to monitor the response of salmon to these events at the CU level.

For decision-making, expert-driven processes will be developed to validate WSP rapid statuses. Expert-driven processes can combine WSP rapid statuses with other information in DFO's Salmon Scanner, supported by expert input and knowledge of CUs. This could be similar to past WSP integrated status assessment processes (Grant & Pestal 2012; Grant et al. 2020; DFO 2015; DFO 2016), using WSP rapid statuses as a foundation. The availability of WSP rapid statuses and associated data can streamline this work going forward, and also support similar Pacific Salmon status assessments conducted by COSEWIC.

One important decision-making context involves applying WSP rapid statuses to the science-based evaluation of limit reference points (LRP) for stock management units (SMUs), which are groups of CUs currently managed together as an aggregate (DFO 2023; Holt et al. 2023a, 2023b). WSP rapid statuses have been recommended as the approach to support this work (DFO 2023; Holt et al. 2023a, 2023b). In LRP evaluations, CU's WSP rapid statuses will be combined with related information provided by CU experts to assign statuses to CUs within SMUs. A *Red* status CU will trigger an SMU rebuilding plan under the *Fisheries Act*. The purpose of using the WSP rapid status approach as a foundation to these SMU LRP assessments is to provide an objective determination of status, grounded in conservation biology principles. This scientific approach also supports standardization and comparability across years and CUs.

The first batch of SMU's prioritized for LRP status assessments include Okanagan Chinook, Interior Fraser coho, and West Coast Vancouver Island Chinook. We are currently working on a standard data summary package and process with DFO, Indigenous and other technical experts related to salmon status and associated data to develop short narratives for the results and concluded outcomes of these assessments. A second DFO technical report (Pestal et al. 2023) provides the framework for the one page results package for each CU, to support the development of narratives.

The WSP rapid statuses with expert input, can also be combined with non-science considerations before and after rebuilding plans are triggered:

- Before rebuilding plans are triggered by DFO science branch, SMUs are prioritized for consideration in the rebuilding plan process. Prioritization includes both scientific and management considerations. Prioritization can include combining WSP rapid statuses with expert input to determine whether or not the SMU is below its LRP. However, prioritization also includes other social, cultural, economic and other factors such as considerations of First Nations Food, Social and Ceremonial needs, international treaty obligations, various stakeholder interests, the vulnerability of CUs to climate change, and more.

- After rebuilding plans are triggered by DFO science using WSP rapid status results and expert input, determination of rebuilding actions is led by management, with scientific inputs. SMU statuses, based on statuses of individual CUs within the LRP process, can be used to help isolate the particular CUs that require rebuilding considerations. This helps to narrow down the scope of the rebuilding plan. It also can help prioritize the type of actions to be taken. For example, though a small but persistent CU may not need specific actions to increase its population size (i.e. rebuild), it likely would require increased protection and maintenance of its existing habitat, due to its small and restricted geographic range and increased extinction risk. CUs of pink or chum salmon, for example, span broad geographic areas in freshwater, therefore, the risk of environmental change or catastrophe are moderated. In contrast, smaller sockeye CUs are likely much more vulnerable to any perturbation or extreme event, which is occurring at an increasing frequency due to climate change.

DFO's Salmon Scanner can be used to integrate knowledge to support climate change vulnerability assessments (CCVA's) and climate change scenario planning. As mentioned, processes will be developed to validate statuses for use within such applications, as determined by experts.

Lastly, WSP rapid status assessments can be used as a prioritization tool for developing WSP integrated statuses using the established process of expert workshops. No official planning cycle has been developed to conduct detailed integrated status assessments or reassessments across the Pacific Region's CUs, even though this was recommended in each of the previous processes (Grant & Pestal 2012; DFO 2015, 2016; Grant et al. 2020). WSP rapid statuses can be used to flag groups of CUs that might require more detailed status assessments. This may be particularly useful for CUs comprised of species or population traits that were not reflected in the previous WSP integrated status assessments. Interpretation and usage of WSP rapid statuses will differ depending on each specific application.

5 CONCLUSIONS

The *Learning Tree 3* algorithm will be used to provide annual WSP rapid statuses for Pacific salmon CUs in BC and the Yukon, with applicable data. This algorithm assigns a *Red*, *Amber* or *Green* status annually, with *High, Medium or Low* confidence ratings. This algorithm performed best across a suite of seven candidate algorithms, when evaluated against quantitative and qualitative criteria. The WSP rapid status algorithm will be broadly accessible to experts through DFO's Salmon Scanner, and interactive data visualization tool to support scientific discovery and decision making.

The WSP rapid status approach ensures that statuses are scientifically objective, consistent, and comparable across BC/Yukon CUs. It also ensures that they are relatively easy to implement, broadly applicable to data rich and data poor CUs, and can be updated annually. This approach is grounded in the principles of conservation biology, which emphasize abundance and trends in abundance criteria to evaluate conservation risk (Caughly 1994;

Mace et al. 2009. They are also grounded in past scientific research and processes where CUs were identified (Holtby & Ciruna 2007; Grant et al. 2011; Brown et al. 2019), and CU statuses were assessed (Holt 2009; Holt et al. 2009; Holt 2010; Grant et al. 2011; Grant and Pestal 2012; DFO 2015, 2016; Grant et al. 2020) (Appendix A).

The WSP rapid status algorithm is designed to be flexible. It can assess status for CUs that have absolute abundance or indices of abundance data. It can improve as more CUs are added for status assessments, and as new methods are developed to consider time-varying productivity in relative abundance benchmarks, *distribution* information, etc. It is named the *Learning Tree* for this reason. If new metrics are added to the algorithm, we recommend that expert-driven WSP integrated status assessment processes are conducted to ground-truth how they influence WSP status determinations.

The ability to track the state and distribution of salmon biodiversity with WSP rapid statuses within DFO's Salmon Scanner comes at a critical time. We are facing a period of accelerating climate and habitat change, which will require timely decisions on where to invest conservation efforts related to salmon and their habitats. The WSP rapid status approach will help support these efforts.

LITERATURE CITED

Brown, G.S., Baillie, S.J., Thiess, M.E., Bailey, R.E., Candy, J.R., Parken, C.K., and Willis, D.M. 2019. Pre-COSEWIC review of southern British Columbia Chinook Salmon (*Oncorhynchus tshawytscha*) conservation units, Part I: background. Can. Sci. Advis. Sec. Res. Doc. 2019/11: vii + 67 pp. Available from https://publications.gc.ca/collections/collection_2019/mpo-dfo/fs70-5/Fs70-5-2019-011-eng.pdf.

Brown, G.S., Holt, C.A., Thiess, M.E., and Pestal, G. 2014. Integrated biological status assessments under the Wild Salmon Policy using standardized metrics and expert judgement: Southern British Columbia Chinook Salmon (*Oncorhynchus tshawytscha*) Conservation Units. Unpublished CSAP Working Paper P67.

Brown, G.S., Thiess, M.E., Wor, C., Holt, C.A., Patten, B., Bailey, R.E., Parken, C.K., Baillie, S.J., Candy, J.R., Willis, D.M., Hertz, E., Connors, B., Pestal, G.P., John, R., Willis, D.M., Hertz, E., Connors, B., and Pestal, G.P. 2020. 2020 Summary of abundance data for Chinook salmon (*Oncorhynchus tshawytscha*) in Southern British Columbia, Canada. Can. Tech. Rep. Fish. Aquat. Sci. 3401: xiii + 214 p. Available from https://waves-vagues.dfo-mpo.gc.ca/Library/40890041.pdf.

Bush, E., and Lemmen, D.S. (Editors). 2019. Canada's changing climate report; Government of Canada, Ottawa, ON. Government of Canada, Ottawa, ON. Available from www.ChangingClimate.ca/CCCR2019.

Caughley, G. 1994. Directions in Conservation Biology. J. Anim. Ecol. 63(2): 215–244. doi:10.2307/5542.

Cheng, L., Abraham, J., Trenberth, K.E., Fasullo, J., Boyer, T., Mann, M.E., Zhu, J., Wang, F., Locarnini, R., Li, Y., Zhang, B., Yu, F., Wan, L., Chen, X., Feng, L., Song, X., and Liu, Y. 2023. Another year of record heat for the oceans. Adv. Atmos. Sci. Available from https://doi.org/10.1007/s00376-023-2385-2.

Cheung, W.W.L., and Frölicher, T.L. 2020. Marine heatwaves exacerbate climate change impacts for fisheries in the northeast Pacific. Sci. Rep. 10(1): 1–10. doi:10.1038/s41598-020-63650-z.

Cheung, W.W.L., Frölicher, T.L., Lam, V.W.Y., Oyinlola, M., Reygondeau, G., Sumaila, U.R., Tai, T.C., Teh, L.C.L., and Wabnitz, C.C.C. 2021. Marine high temperature extremes amplify the impacts of climate change on fish and fisheries. Sci. Adv.: 1–16. doi:https://www.science.

COSEWIC. 2010. COSEWIC's assessment process and criteria. Available from https://www.canada.ca/content/dam/eccc/migration/cosewic-cosepac/94d0444d-369c-49ed-a586-ec00c3fef69b/assessment_process_and_criteria_e.pdf.

COSEWIC. 2020. COSEWIC assessment and status report on the Chinook Salmon *Oncorhynchus tshawytscha*, Designatable Units in Southern British Columbia (Part One – Designatable Units with no or low levels of artificial releases in the last 12 years), in Canada. : xxxi + 283 pp. Available from https://wildlife-species.canada.ca/species-risk-registry/virtual_sara/files/cosewic/ChinookSalmon-v00-2019-Eng.pdf.

Crozier, L.G., Burke, B.J., Chasco, B.E., Widener, D.L., and Zabel, R.W. 2021. Climate change threatens Chinook salmon throughout their life cycle. Commun. Biol. 4(1): 222.

Springer US. doi:10.1038/s42003-021-01734-w.

Crozier, L.G., McClure, M.M., Beechie, T., Bograd, S.J., Boughton, D.A., Carr, M., Cooney, T.D., Dunham, J.B., Greene, C.M., Haltuch, M.A., Hazen, E.L., Holzer, D.M., Huff, D.D., Johnson, R.C., Jordan, C.E., Kaplan, I.C., Lindley, S.T., Mantua, N.J., Moyle, P.B., Myers, J.M., Nelson, M.W., Spence, B.C., Weitkamp, L.A., Williams, T.H., and Willis-Norton, E. 2019. Climate vulnerability assessment for Pacific salmon and steelhead in the California Current Large Marine Ecosystem. PLoS One 14(7): e0217711. doi:10.1371/journal.pone.0217711.

DFO. 2005. Canada's Policy for Conservation of Wild Pacific Salmon. Fisheries and Oceans Canada, Vancouver, B.C., pp. vi+ 49. Available from https://www.pac.dfo-mpo.gc.ca/fm-gp/species-especes/salmon-saumon/wsp-pss/policy-politique/index-eng.html.

DFO. 2012. Integrated biological status of Fraser River sockeye salmon (*Oncorhynchus nerka*) under the Wild Salmon Policy. Can. Sci. Advis. Sec. Sci. Advis. Rep. 2012/056: 13 pp. Available from http://www.dfo-mpo.gc.ca/csas-sccs/Publications/SAR-AS/2012/2012_056-eng.html.

DFO. 2015. Wild Salmon Policy status assessment for conservation units of Interior Fraser River coho (*Oncorhynchus kisutch*). Can. Sci. Advis. Sec. Sci. Advis. Rep. 2015/022: 12 pp. Available from https://waves-vagues.dfo-mpo.gc.ca/Library/364851.pdf.

DFO. 2016. Integrated biological status of southern British Columbia Chinook salmon (*Oncorhynchus tshawytscha*) under the Wild Salmon Policy. Can. Sci. Advis. Sec. Sci. Advis. Rep. 2016/042: 15 pp. Available from https://waves-vagues.dfo-mpo.gc.ca/Library/40595419.pdf.

DFO. 2018. The 2017 Fraser Sockeye salmon (*Oncorhynchus nerka*) integrated biological status re-assessments under the Wild Salmon Policy. Can. Sci. Advis. Sec. Sci. Advis. Rep. 2018/017: 17 pp. Available from http://waves-vagues.dfo-mpo.gc.ca/Library/40712163.pdf.

DFO. 2023. Biological benchmarks and building blocks for aggregate-level management targets for Skeena and Nass Sockeye salmon (*Oncorhynchus nerka*). Can. Sci. Advis. Sec. Sci. Advis. Rep. 2023/008. pp. 20. Available from https://www.dfo-mpo.gc.ca/csas-sccs/Publications/SAR-AS/2023/2023_008-eng.pdf.

Dorner, B., Catalano, M.J., and Peterman, R.M. 2018. Spatial and temporal patterns of covariation in productivity of Chinook salmon populations of the northeastern Pacific Ocean. Can. J. Fish. Aquat. Sci. 75(7): 1082–1095. doi:10.1139/cjfas-2017-0197.

Dorner, B., Peterman, R.M., and Haeseker, S.L. 2008. Historical trends in productivity of 120 Pacific pink, chum, and sockeye salmon stocks reconstructed by using a Kalman filter. Can. J. Fish. Aquat. Sci. 65(9): 1842–1866. doi:https://doi.org/10.1139/F08-094.

Government of B.C., DFO, and FRAFS. 2019, September 8. Salmon swimming past Big Bar. Information Bulletin prepared by the Government of B.C., Fisheries and Oceans Canada, and the Fraser River Aboriginal Fisheries Secretariat. Available from https://www2.gov.bc.ca/assets/gov/public-safety-and-emergency-services/emergency-preparedness-response-recovery/embc/big-bar-landslide-2019/19_71w20ay_information_bulletin_-_fish_passage.pdf.

Grant, S.C.H., Holt, C.A., Pestal, G., Davis, B.M., and MacDonald, B.L. 2020. The 2017 Fraser Sockeye salmon (*Oncorhynchus nerka*) integrated biological status re-assessments under the Wild Salmon Policy using standardized metrics and expert judgement. Can. Sci. Advis. Sec. Res. Doc. 2020/038: vii+ 211. Available from http://www.dfo-mpo.gc.ca/csas-sccs/Publications/ResDocs-DocRech/2020/2020_038-eng.pdf.

Grant, S.C.H., MacDonald, B.L., Cone, T.E., Holt, C.A., Cass, A., Porszt, E.J., Hume, J.M.B., and Pon, L.B. 2011. Evaluation of uncertainty in Fraser sockeye (*Oncorhynchus nerka*) Wild Salmon Policy status using abundance and trends in abundance metrics. Can. Sci. Advis. Sec. Res. Doc. 2011/087: viii + 183 pp. Available from https://science-catalogue.canada.ca/record=4054219~S6.

Grant, S.C.H., MacDonald, B.L., Lewis, D., G.J., N.L.W.C., and Michielsens, C.G.J. 2021. State of Canadian Pacific salmon in 2020. In State of the Physical, Biological and Selected Fishery Resources of Pacific Canadian Marine Ecosystems in 2020. Can. Tech. Rep. Fish. & Aquat. Sci. 3434. pp. vii + 231. Edited by J.L. Boldt, A. Javorski, and P.C. Chandler.

Grant, S.C.H., MacDonald, B.L., and Winston, M.L. 2019. State of the Canadian Pacific salmon: responses to changing climate and habitats. Can. Tech. Rep. Fish. Aquat. Sci. 3332: ix + 50 pp. Available from http://www.dfo-mpo.gc.ca/species-especes/publications/salmon-saumon/state-etat-2019/abstract-resume/index-eng.html.

Grant, S.C.H., and Pestal, G. 2012. Integrated biological status assessments under the Wild Salmon Policy using standardized metrics and expert judgement: Fraser River sockeye salmon (*Oncorhynchus nerka*) case studies. Can. Sci. Advis. Sec. Res. Doc. 2012/106: v + 132 pp. Available from https://waves-vagues.dfo-mpo.gc.ca/Library/349637.pdf.

Hare, J.A., Morrison, W.E., Nelson, M.W., Stachura, M.M., Teeters, E.J., Griffis, R.B., Alexander, M.A., Scott, J.D., Alade, L., Bell, R.J., Chute, A.S., Curti, K.L., Curtis, T.H., Kircheis, D., Kocik, J.F., Lucey, S.M., McCandless, C.T., Milke, L.M., Richardson, D.E., Robillard, E., Walsh, H.J., McManus, M.C., Marancik, K.E., and Griswold, C.A. 2016. A vulnerability assessment of fish and invertebrates to climate change on the Northeast U.S. continental shelf. PLoS One 11(2): 30 pp. doi:10.1371/journal.pone.0146756.

Hayes, S.A., and Kocik, J.F. 2014. Comparative estuarine and marine migration ecology of Atlantic salmon and steelhead: Blue highways and open plains. Rev. Fish Biol. Fish. 24(3): 757–780. doi:10.1007/s11160-014-9348-8.

Holt, C., Davis, B., Dobson, D., Godbout, L., Luedke, W., Tadey, J., and Van Will, P. 2018. Evaluating benchmarks of biological status for data-limited conservation units of Pacific salmon, focusing on chum salmon in Southern BC. DFO Can. Sci. Advis. Sec. Res. Doc. (2018/011): ix + 77. Available from http://www.dfo-mpo.gc.ca/csas-sccs/csas-sccs@dfo-mpo.gc.ca.

Holt, C.A. 2009. Evaluation of benchmarks for conservation units in Canada's Wild Salmon Policy: technical documentation. Can. Sci. Advis. Sec. Res. Doc. 2009/059: x + 50. Available from https://www.dfo-mpo.gc.ca/csas-sccs/publications/resdocs-docrech/2009/2009_059-eng.htm.

Holt, C.A. 2010. Will depleted populations of Pacific salmon recover under persistent reductions in survival and catastrophic mortality events? ICES J. Mar. Sci. 67(9): 2018–2026. Available from https://academic.oup.com/icesjms/article/67/9/2018/620513.

Holt, C.A., Cass, A., Holtby, B., and Riddell, B. 2009. Indicators of status and benchmarks for Conservation Units in Canada's Wild Salmon Policy. Can. Sci. Advis. Sec. Res. Doc. 2009/058: viii + 74. Available from https://www.dfo-mpo.gc.ca/csas-sccs/publications/resdocs-docrech/2009/2009_058-eng.htm.

Holt, C.A., Holt, K., Wor, C., Warkentin, L., Connors, B., Grant, S.C.H., and Huang, A.-M. 2023a. Guidelines for Defining Limit Reference Points for Pacific Salmon Stock Management Units. Can. Sci. Adv. Sec. Res. Doc. **2023/009**: iv + 66. Available from https://www.dfo-mpo.gc.ca/csas-sccs/Publications/ResDocs-DocRech/2023/2023_009-eng.html.

Holt, K., Holt, C.A., Warkentin, L., Wor, C., Davis, B., Arbeider, M., Bokvist, J., Crowley, S., Grant, S.C.H., Luedke, W., McHugh, D., Picco, C., and Will, P. Van. 2023b. Case Study Applications of LRP Estimation Methods to Pacific Salmon Stock Management Units. Can. Sci. Advis. Sec. Res. Doc. **2023/010**: iv + 129. Available from https://www.dfo-mpo.gc.ca/csas-sccs/Publications/ResDocs-DocRech/2023/2023_010-eng.html.

Holtby, B.L., and Ciruna, K.A. 2007. Conservation units for Pacific salmon under the Wild Salmon Policy. Can. Sci. Advis. Sec. Res. Doc. 2007/070: viii + 350 pp. Available from http://www.dfo-mpo.gc.ca/csas-sccs/publications/resdocs-docrech/2007/2007_070-eng.htm.

IPCC. 2021. Summary for policymakers. *In Climate Change 2021: The Physical Science Basis. Contribution of Working Group 1 to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*. Edited by V. Masson-Delmotte, P. Zhai, A. Pirani, S.L. Connors, C. Péan, S. Berger, N. Caud, Y. Chen, L. Goldfarb, M.I. Gomis, M. Huang, K. Leitzell, E. Lonnoy, J.B.R. Matthews, T.K. Maycock, T. Waterfield, O. Yelekçi, R. Yu, and B. Zhou. Cambridge University Press, Cambridge, UK and New York, NY, USA. pp. 3–22. doi:10.1017/CBO9781139177245.003.

IPCC. 2022a. Climate Change 2022: Impacts, Adaptation and Vulnerability Summary for Policy Makers. doi:10.4324/9781315071961-11.

IPCC. 2022b. Climate Change 2022: Impacts, Adaptation and Vulnerability Report. doi:10.4324/9781315071961-11.

Irvine, J.R., Chapman, K., and Park, J. 2019. Report of the proceedings for the IYS workshop. International Year of the Salmon workshop on salmon status and trends. North Pacific Anadromous Fish Comm. Tech. Rep. 13. Vancouver, B.C. 99 pp. Available from https://npafc.org/wp-content/uploads/2019/08/Tech-Rep-13_Final_16Aug2019.pdf.

IUCN. 2022. Guidelines for Using the IUCN Red List Categories and Criteria. Available from https://www.iucnredlist.org/resources/redlistguidelines.

MacDonald, B.L., and Grant, S.C.H. 2012. Pre-season run size forecasts for Fraser River sockeye salmon (*Oncorhynchus nerka*) in 2012. Can. Sci. Advis. Sec. Res. Doc. 2012/011: v + 64 pp. Available from http://www.dfo-mpo.gc.ca/csas-sccs/Publications/ResDocs-DocRech/2012/2012_145-eng.html.

MacDonald, B.L., Grant, S.C.H., and Wilson, N.L. 2023. State of Canadian Pacific salmon in 2022. *In State of the Physical, Biological and Selected Fishery Resources of Pacific Canadian Marine Ecosystems in 2022*. Can. Tech. Rep. Fish. Aquat. Sci. 3542. pp. 125–132. Available from https://waves-vagues.dfo-mpo.gc.ca/library-bibliotheque/41199248.pdf.

Mace, G., Collar, J., Cooke, J.G., Gaston, K.J., Ginsberg, N., Leader Williams, N., Maunder, M., and Milner-Gulland, E.J. 1992. The development of new criteria for listing species on the IUCN Red List. 19: 16–22. Available from https://www.researchgate.net/publication/236679197_The_development_of_new_criteria_for_listing_species_on_the_IUCN_Red_List.

Mace, G.M., Collar, N.J., Gaston, K.J., Hilton-Taylor, C., Arçakaya, H.R., Leader-Williams, N., and Stuart, S.N. 2008. Quantification of extinction risk : IUCN's system for classifying threatened species. Conserv. Biol. 22(6): 1424–1442. doi:10.1111/j.1523-1739.2008.01044.x.

Mace, G.M., and Lande, R. 1991. Assessing extinction threats : toward a reevaluation of IUCN threatened species categories. Conserv. Biol. 5(2): 148–157. doi:10.1111/j.1523-1739.1991.tb00119.x.

Malick, M.J., Cox, S.P., Mueter, F.J., Dorner, B., and Peterman, R.M. 2017. Effects of the North Pacific Current on the productivity of 163 Pacific salmon stocks. Fish. Oceanogr. 26(3): 268–281. doi:10.1111/fog.12190.

McElhany, P., Ruckelshaus, M.H., Ford, M.J., Wainwright, T.C., and Bjorkstedt, E.P. 2000. Viable salmonid populations and the recovery of evolutionarily significant units. U.S. Dept. Commer., NOAA Tech. Memo. NMFS-NWFSC-42. pp. 156. Available from http://www.orafs.org/pdfdocs/Comment-2001 UWRCFMEP.pdf.

National Research Council (US) Committee on Scientific Issues in the Endangered Species Act. 1998. Science and the Endangered Species Act. doi:10.2307/5747.

Parken, C.K., McNicol, R.E., and Irvine, J.R. 2006. Habitat-based methods to estimate escapement goals for data limited Chinook salmon stocks in British Columbia, 2004. Can. Sci. Advis. Secr. Res. Doc. 2006/083: vii + 74 pp. Available from https://waves-vagues.dfo-mpo.gc.ca/library-bibliotheque/326898.pdf.

Parken, C.K., Ritchie, L., Macdonald, B.L., Bailey, R.E., Nicklin, P., Bradford, M.J., Ward, H., Welch, P., Boyce, I., Tompkins, A., Maxwell, M., Beach, K., Irvine, J.R., Grant, S.C.H., Van Will, P., Willis, D., Staley, M.J., Walsh, M., Sawada, J., Scroggie, J., and McGrath, E. 2014. Wild Salmon Policy biological status assessment for conservation units of Interior Fraser River coho salmon (*Oncorhynchus kisutch*). Unpublished CSAP Working Paper 2013SAL12.

Pestal, G., MacDonald, B.L., Grant, S.C.H., and Carr-Harris, C. 2023. State of the Salmon: Wild Salmon Policy rapid status approach for Fraser, Skeena, and Nass sockeye, Fraser pink and chum, Interior Fraser Coho, and southern BC Chinook salmon. Can. Tech. Rep. Fish. Aquat. Sci. (in review).

Peterman, R.M., Pyper, B.J., and MacGregor, B.W. 2003. Use of the Kalman filter to reconstruct historical trends in productivity of Bristol Bay sockeye salmon (*Oncorhynchus nerka*). Can. J. Fish. Aquat. Sci. 60: 809–824. doi:10.1139/F03-069.

Picard, R.R., and Cook, R.D. 1984. Cross-validation of regression models. J. Am. Stat. Assoc. 79(387): 575–583. doi:10.1080/01621459.1984.10478083.

Ripley, B.D. 1996. Pattern Recognition and Neural Networks. Cambridge University Press, Cambridge, U.K.

Rodrigues, A.S.L., Pilgrim, J.D., Lamoreux, J.F., Hoffmann, M., and Brooks, T.M. 2006. The value of the IUCN Red List for conservation. Trends Ecol. Evol. 21(2): 71–76. doi:10.1016/j.tree.2005.10.010.

Schindler, D.E., Hilborn, R., Chasco, B., Boatright, C.P., Quinn, T.P., Rogers, L.A., and Webster, M.S. 2010. Population diversity and the portfolio effect in an exploited species. Nature 465(7298): 609–612. doi:10.1038/nature09060.

Therneau, T., and Atkinson, B. 2019. Rpart: recursive partitioning and regression trees.

Available from https://cran.r-project.org/web/packages/rpart/rpart.pdf.

Wade, J., Hamilton, S., Baxter, B., Brown, G., Grant, S.C.H., Holt, C.A., Thiess, M., and Withler, R.E. 2019. Framework for reviewing and approving revisions to Wild Salmon Policy Conservation Units. Can. Sci. Advis. Sec. Res. Doc. 2019/015: v + 29 pp. Available from http://waves-vagues.dfo-mpo.gc.ca/Library/40780399.pdf.

Withler, R.E., Bradford, M.J., Willis, D., and Holt, C.A. 2018. Genetically Based Targets for Enhanced Contributions to Canadian Pacific Chinook Salmon Populations. DFO Can. Sci. Advis. Sec. Sci. Advis. Rep. 2018/019(001): 88 p. Available from https://waves-vagues.dfo-mpo.gc.ca/Library/40702285.pdf.

TABLES

Table 1: **Biological status zones under the Wild Salmon Policy (WSP).** Note that although the WSP initially only identified three status zones (DFO 2005), two additional status zones were added in subsequent WSP integrated assessment processes: *Red/Amber* and *Amber/Green* (Grant & Pestal 2012). WSP integrated status assessments also designated some CUs as either data deficient (DD) or undetermined (UD). As part of the current work on WSP rapid status assessments, we also tested a lower-resolution status scale with 2 zones: *Red* vs. *Not Red* (see Table 4).

Status		Definition
	Red	Poor status CU at imminent threat of extinction [revised definition, given alignment with COSEWIC *Endangered* status zone]
	Red/ Amber	Intermediate between *Red* and *Amber*
	Amber	"While a CU in the *Amber* zone should be at low risk of loss, there will be a degree of lost production. Still, this situation may results when CUs share risk factors with other, more productive units"
	Amber/ Green	Intermediate between *Amber* and *Amber*
	Green	"identif[ies] whether harvest are greater than the level expected to provide on an average annual basis, the maximum annual catch for a CU, given existing conditions…there would not be a high probability of losing the CU"
	DD	**Data deficient.** CUs have been designated as DD if there is no data available, or if the available data is insufficient for calculating status metrics (after quality control).
	UD	**Undetermined.** CUs have been designated as UD if data are available and metrics have been calculated, but the expert participants in the status workshops could not settle on a consensus WSP integrated status.
	Not Red	Used explicitly in the 'Simply *Red*' algorithm and on a 2 status scale: *Red* versus *Not Red* (see **Table 4**). *Not Red* includes: *Amber, Amber/Green*, and *Green* statuses. Conversely, the *Red* status zone includes: *Red* and *Red/Amber* statuses.

Table 2: **Classification Tree Terminology.** The approach of extracting algorithms using *Classification and Regression Trees* (CART) is widely used in decision analysis and machine learning. The approach is very versatile but comes with highly specialized terminology. This table defines key terms.

Term	Description
Classification Tree	Nested sequence of criteria that classifies cases into different categories (e.g. field guide for species identification).
Binary Split	Criterion for separating a set of cases into 2 subsets
Recursive Partitioning	Step-wise splitting of a sample into smaller subsets of cases that are similar to each other (e.g. sort test fishing catch by species, then by sex, then by size category).
Node	A fork in the classification tree. Nodes are systematically numbered. The root node is 1. Nodes created by each binary split are numbered as double for the NO subset and double +1 for the YES subset (e.g. node 4 splits into nodes 8 and 9). This way, even-numbered nodes are always the result of NO in the parent node, and the number of each node uniquely defines the full path up to that node (e.g. cases in node 17 are the result of YES in node 8, NO in node 4, NO in node 2, and NO in node 1).
Branch	A node the leads to another node
Leaf (a.k.a terminal node)	A node at the end of a branch, resulting in a classification.
Loss function	Assigns a penalty for each type of error. This is a more flexible version of the sum-of-squared errors in a regression fit, which can be customized to handle qualitative errors (e.g. species misidentifications).
Asymmetric loss function	Classification errors may have different implications depending on the direction of the error (e.g. classifying a poisonous mushroom as edible has worse consequences than the reverse error). These considerations can be built into the tree fitting step by specifying an asymmetric loss function, that results in heavier penalties for one type of error compared to another.
Complexity Penalty	Tree fitting can apply a penalty for the number of branches, to 'prune' the tree and avoid overfitting the data. This is equivalent to choosing among alternative regression models based on criteria that consider model complexity, such as adjusted R^2 or *Akaike's Information Criterion* (AIC).
Learning Data Set	Sample of cases where the correct classification is known. This is used to fit the trees and is equivalent to the data used to estimate parameters for a forecasting model.
Test Data Set	New set of cases (with or without known correct classifications) used to assess performance of the fitted tree. This is equivalent to generating a forecast based on estimated parameters and new values for the predictor variables.
Confusion Matrix	2-way contingency table showing true classifications versus classifications assigned by the tree.
Surrogate Split	Software used to fit classification trees also identifies surrogate splitting criteria. If data for the main criterion is not available, it will use any alternate information that closely replicates the split generated by the main criterion.

Table 3: **Alternative Settings for CART Explorations**. When fitting classification trees, similar to fitting regression models, there are many available options for the inputs, outputs, and settings (e.g. variables to include, variable transformations, alternative model forms). This table lists the variations we explored and screened to develop a shortlist of *fitted* algorithms for more detailed evaluation (Section 2.4.2). R: *Red;* RA: *Red/Amber; A: Amber;* AG: *Amber/Green;* G: *Green;* DD: *data deficient;* UD: *undetermined.*

Setting	Variations
Response Variable	• WSP integrated statuses (R/RA/A/AG/G) • simplified statuses (R/A/G) • WSP integrated status scores (1-5) • simplified status scores (1-3) • exponential WSP integrated status scores (e.g. G=1,AG=2,A=4,RA=8, R=16,DD=0,UD=0)
Predictor Variable	• Numeric metric values • metric statuses (*R/A/G*)
Model Fits	• complexity penalty • asymmetric penalty function (error direction matters)
Data Set	• use all cases • fit separate tree by species • fit separate trees by data type

Table 4: **Alternative status scales for evaluating algorithm performance.** WSP rapid statuses were converted to scores from 1 to 5 to capture the magnitude and direction of classification errors. WSP rapid statuses need to be simplified to the same scale as the WSP integrated status assessments to make meaningful comparisons within algorithm and between algorithms. The five status scale is the scale used in previous WSP integrated status assessments. Different WSP rapid status candidate algorithms were developed to assess status on one of the three scales (Table 6), with statuses converted to each of the three scales for performance evaluation (Table 7).

	5 Status Scale		**3 Status Scale**		**2 Status Scale**	
	Zone	**Score**	**Zone**	**Score**	**Zone**	**Score**
	Red	5	*Red*	5	*Red*	5
	Red/Amber	4				
	Amber	3	*Amber*	3	*Not Red*	2
	Amber/Green	2				
	Green	1	*Green*	1		

Table 5: **Total completed WSP integrated status assessments for Fraser sockeye, Southern BC Chinook, Interior Fraser coho.** This table shows the frequency of status designations from the three integrated assessments and one re-assessment for Fraser sockeye completed under the WSP (Grant and Pestal 2012; DFO 2015, 2016). To create more comparable results with particular algorithms, the five status scale was converted into the three (*Red*, *Amber* or *Amber*), or two (*Red/Not Red*) status scales (see Table 4 above). The table also summaries the number of Data Deficient (DD) CUs and Undetermined (UD) statuses for each species. Detailed metrics and statuses in Appendix C. *Total CUs Assessed* is less than *Total Statuses Assigned,* because Fraser sockeye have been assessed twice (22 CUs in first assessment, 23 CUs in the reassessment).

Status Zone	**Sockeye Status Scale**			**Chinook Status Scale**			**Coho Status Scale**			**Total Status Scale**		
	5	3	2	5	3	2	5	3	2	5	3	2
Red	14	20	20	11	12	12	0	0	0	25	32	32
Red/Amber	6			1			0			7		
Amber	9	17	25	1	1	3	3	5	5	13	23	33
Amber/Green	8			0			2			10		
Amber	8	8		2	2		0	0		10	10	
DD	0	0	0	4	4	4	0	0	0	4	4	4
UD	1	1	1	0	0	0	0	0	0	1	1	1
Total CUs Assessed	**23**	**23**	**23**	**19**	**19**	**19**	**5**	**5**	**5**	**47**	**47**	**47**
Total Statuses Assigned	**45**	**45**	**45**	**15**	**15**	**15**	**5**	**5**	**5**	**65**	**65**	**65**
Red Status	0.31	0.44	0.44	0.73	0.8	0.8	0	0	0	0.38	0.49	0.49

Table 6**: The seven candidate rapid status algorithms. Three fitted algorithms** based on exploring alternative CART model fits, and **four constructed algorithms** based on combining CART fits with additional considerations. The fitted algorithms used all 65 cases from the *learning data set*. Exploratory CART fits using data split by species or data type were unstable and were therefore not included in the shortlist of candidate algorithms for detailed testing. Constructed algorithms use components of the fitted algorithms, so indirectly use all available data. This table summarizes the design approach for each algorithm. Section 2.4 describes the development steps. Appendix D shows the full algorithms as a diagram and as a set of classification rules. Note that these algorithms generate rapid statuses at different scales of resolution, from 5 (*Red, Red/Amber, Amber, Amber/Green, Green*), to 3 (*Red, Amber, Green*) to 2 (*Red, Not Red*), as shown by the 'x' in the right-hand columns (R = *Red*, nR = *Not Red*, RA = *Red/Amber*, A = *Amber*, AG = *Amber/Green*, G = *Green*) (see Table 4).

Type	Name	Description	R	nR	RA	A	AG	G
Fitted	**Minimalist** Appendix E.1	• 3 status scale: simplified status scale • Built using only the values for trend metrics: *long-term & percent change*, which are broadly available metrics common to most CUs • Tree fitting with high complexity penalty to generate a simple tree with few branches.	X			X		X
	Fancy Pants Appendix E.2	• 5 status scale: matches WSP integrated status scale • Built using values for all available metrics • Tree fitting with low complexity penalty to generate a more complex tree with finer resolution with more branches.	X		X	X	X	X
	Categorical Realist Appendix E.3	• 2 status scale: simplified status scale • Simplified metrics: *absolute abundance*, *relative abundance* and *long-term trend* • Fit separate trees for different data types, but only R and A were isolated as terminal nodes by the tree fit.	X			X		
Constr-ucted	**Simply Red** Appendix E.4	• 2 status scale: simplified status scale • Simplified metrics: *long-term trend, percent change*, and *relative abundance* • Combines all the criteria from the other algorithms that flag a R status	X	X				
	Learning Tree 1 Appendix E.5	• 3 status scale: simplified status scale • Built on the CART algorithms but combined with WSP integrated status assessment narratives.	X			X		X
	Learning Tree 2 Appendix E.6	• 3 status scale: simplified status scale • Same as *Learning Tree 1* but use R/A/G metrics instead of metric values.	X			X		X
	Learning Tree 3 Appendix E.7	• 3 status scale: simplified status scale • Evolution of *Learning Tree 1*, putting *absolute abundance* first, and providing additional considerations for long-term trend and percent change metrics.	X			X		X

Table 7: **Summary of algorithm performance across all 65 cases in the *learning data set*: Fraser sockeye, Southern BC Chinook and Interior Fraser coho CUs.** Note in the *learning set* there are two years with WSP integrated status assessments completed for Fraser sockeye CUs, in addition to one year for Southern BC Chinook CUs and one year for Interior Fraser coho CUs. The table shows the completion rate (*Number Complete*): the number of cases the algorithm could assign a WSP rapid status to out of the total 65 *learning data set* cases; number of correct designations (*Number Correct):* the rapid status matches WSP integrated status; the number of close designations (*Number Close*): the rapid status is only 1 status zone different from the WSP integrated status; and the number of overestimates (*Number Predicted Better):* the rapid status is better than the WSP integrated status. *Median*, *Mean* and *Range of Errors* are presented in the last 3 rows**.** All errors are calculated by converting status designations to a 2, 3, or 5-status scale (see Table 4 to Table 6 for details). The status scale that matches the algorithm is marked with bold font and grey shading. A negative error means that the algorithm assigned a worse status than the integrated expert assessment. Table cells are highlighted in orange if a rapid status could be assigned for less than 3/4 of the cases (*Number Complete* < 49/65), or if the mean error was larger than 0.3 (Mean < -0.3 or Mean > 0.3).
Table 8 to Table 10 show the same summary by species.

Performance Measure	Status Scale	Minimalist	Fancy Pants	Categorical Realist	Simply Red	Learning Tree 1	Learning Tree 2	Learning Tree 3
Number Complete		64	54	55	55	65	65	65
Number Correct	5	39	**47**	30	23	39	41	44
	3	**49**	49	**41**	26	**46**	**48**	**54**
	2	55	50	50	**47**	58	58	59
Number Close	5	54	**50**	44	47	55	58	60
	3	**49**	49	**41**	47	**46**	**48**	**54**
	2	55	50	50	**47**	58	58	59
Number Predicted Better	5	8	**2**	5	10	17	16	7
	3	**8**	1	**5**	17	**17**	**16**	**7**
	2	6	0	5	**2**	6	7	4
Median Error	5	0	**0**	0	0	0	0	0
	3	**0**	0	**0**	0	**0**	**0**	**0**
	2	0	0	0	**0**	0	0	0
Mean Error	5	0.25	**0.2**	0.4	0.27	-0.23	-0.23	0.2
	3	**0**	0.19	**0.15**	0.02	**-0.49**	**-0.49**	**-0.06**
	2	-0.14	0.22	-0.27	**0.22**	-0.23	-0.32	-0.09
Range of Error	5	-2 to 4	**-1 to 4**	-2 to 2	-2 to 4	-4 to 4	-4 to 2	-2 to 4
	3	**-2 to 4**	-2 to 4	**-2 to 2**	-3 to 4	**-4 to 4**	**-4 to 2**	**-2 to 4**
	2	-3 to 4	0 to 4	-3 to 2	**-3 to 4**	-3 to 4	-3 to 2	-3 to 4

Table 8: **Fraser sockeye summary of algorithm performance on the *learning data set*.** Layout as per Table 7. There are a total of 45 cases in the *learning data set* for Fraser sockeye (see Table 5).

Performance Measure	Status Scale	Minimalist	Fancy Pants	Categorical Realist	Simply Red	Learning Tree 1	Learning Tree 2	Learning Tree 3
Number Complete	NA	44	38	35	39	45	45	45
Number Correct	5	25	**31**	17	13	27	29	29
	3	**33**	33	**26**	16	**34**	**36**	**38**
	2	37	34	33	**31**	40	41	41
Number Close	5	37	**34**	28	31	41	43	42
	3	**33**	33	**26**	31	**34**	**36**	**38**
	2	37	34	33	**31**	40	41	41
Number Predicted Better	5	6	**2**	2	7	9	8	4
	3	**6**	1	**2**	12	**9**	**8**	**4**
	2	4	0	2	**2**	4	4	2
Median	5	0	**0**	0	0	0	0	0
	3	**0**	0	**0**	0	**0**	**0**	**0**
	2	0	0	0	**0**	0	0	0
Mean	5	0.3	**0.29**	0.6	0.44	0.04	0	0.31
	3	**0**	0.26	**0.29**	0.13	**-0.27**	**-0.31**	**0**
	2	-0.07	0.32	-0.17	**0.31**	-0.2	-0.27	0
Range	5	-2 to 4	**-1 to 4**	-1 to 2	-2 to 4	-2 to 4	-2 to 2	-1 to 4
	3	**-2 to 4**	-2 to 4	**-2 to 2**	-3 to 4	**-2 to 4**	**-2 to 2**	**-2 to 4**
	2	-3 to 4	0 to 4	-3 to 2	**-3 to 4**	-3 to 4	-3 to 2	-3 to 4

Table 9: **Southern BC Chinook summary of algorithm performance in the *learning data set*.** Layout as per Table 7. There are a total of 15 cases in the Learning data set for Southern BC Chinook (see Table 5).

Performance Measure	Status Scale	Minimalist	Fancy Pants	Categorical Realist	Simply Red	Learning Tree 1	Learning Tree 2	Learning Tree 3
Number Complete	NA	15	11	15	11	15	15	15
Number Correct	5	11	**11**	10	10	12	12	12
	3	**11**	11	**10**	10	**12**	**12**	**12**
	2	13	11	12	**11**	13	12	13
Number Close	5	12	**11**	11	11	12	13	13
	3	**11**	11	**10**	11	**12**	**12**	**12**
	2	13	11	12	**11**	13	12	13
Number Predicted Better	5	2	**0**	3	0	3	3	2
	3	**2**	0	**3**	0	**3**	**3**	**2**
	2	2	0	3	**0**	2	3	2
Median	5	0	**0**	0	0	0	0	0
	3	**0**	0	**0**	0	**0**	**0**	**0**
	2	0	0	0	**0**	0	0	0
Mean	5	0.07	**0**	-0.07	0.09	-0.6	-0.47	-0.07
	3	**0**	0	**-0.13**	0.09	**-0.67**	**-0.53**	**-0.13**
	2	-0.4	0	-0.6	**0**	-0.4	-0.6	-0.4
Range	5	-2 to 2	**0 to 0**	-2 to 2	0 to 1	-4 to 0	-4 to 0	-2 to 2
	3	**-2 to 2**	0 to 0	**-2 to 2**	0 to 1	**-4 to 0**	**-4 to 0**	**-2 to 2**
	2	-3 to 2	0 to 0	-3 to 2	**0 to 1**	-3 to 0	-3 to 0	-3 to 2

Table 10: **Interior Fraser coho summary of algorithm performance in the *learning data set*.** Layout as per Table 7. There are a total of 5 cases in the *learning data set* for Interior Fraser coho (see Table 5).

Performance Measure	Status Scale	Minimalist	Fancy Pants	Categorical Realist	Simply Red	Learning Tree 1	Learning Tree 2	Learning Tree 3
Number Complete	NA	5	5	5	5	5	5	5
Number Correct	5	3	**5**	3	0	0	0	3
	3	**5**	5	**5**	0	**0**	**0**	**4**
	2	5	5	5	**5**	5	5	5
Number Close	5	5	**5**	5	5	2	2	5
	3	**5**	5	**5**	5	**0**	**0**	**4**
	2	5	5	5	**5**	5	5	5
Number Predicted Better	5	0	**0**	0	3	5	5	1
	3	**0**	0	**0**	5	**5**	**5**	**1**
	2	0	0	0	**0**	0	0	0
Median	5	0	**0**	0	-1	-2	-2	0
	3	**0**	0	**0**	-1	**-2**	**-2**	**0**
	2	0	0	0	**0**	0	0	0
Mean	5	0.4	**0**	0.4	-0.6	-1.6	-1.6	0
	3	**0**	0	**0**	-1	**-2**	**-2**	**-0.4**
	2	0	0	0	**0**	0	0	0
Range	5	0 to 1	**0 to 0**	0 to 1	-1 to 0	-2 to -1	-2 to -1	-1 to 1
	3	**0 to 1**	0 to 0	**0 to 1**	-1 to 0	**-2 to -1**	**-2 to -1**	**-2 to 1**
	2	0 to 1	0 to 0	0 to 1	**0 to 0**	0 to -1	0 to -1	0 to 1

Table 11: **Summary of the *relative abundance* metric sensitivity test that compares how rapid statuses change when this metric was included or excluded from a CU's metric set.** The *relative abundance* metric is available for 37 of the 65 cases in the *learning data set*. This metric is used by the following algorithms: *Categorical Realist*, *Simply Red*, and *Learning Tree 1, 2*, and *3*. The *Minimalist* does not use the *relative abundance* metric, so it was excluded. The *Categorical Realist* and *Simply Red* were included, but since they are relatively simple, they cannot assign statuses for 37/37 and 25/37 cases, respectively. This table shows the number of cases where the algorithm could assign status (Number Completed) vs. where the algorithm could not assign a status (Number Not Completed). This is presented for the two scenarios: With and Without *relative abundance* metric (RA). It also shows the number of cases where the rapid status changed, as well as the direction and magnitude of the changes. Notable results are highlighted in orange. The asterisks denotes where excluding the *relative abundance* metric results in an incomplete status for Chilliwack-ES sockeye, which is an exceptional case (see Section 3.5 for description of the Chilliwack-ES sockeye CU exception).

		Fitted	***Constructed***			
				Learning Tree		
	Measure	Categorical Realist	Simply *Red*	LT1	LT2	LT3
With RA metric	Number Completed	37	36	37	37	37
	Number Not Completed	0	1	0	0	0
Without RA metric	Number Completed	0	12	36	36	36
	Number Not Completed	37	25	1*	1*	1*
	Number Changed	0	0	17	13	9
	Number Worse by 1 status zone	0	0	2	5	5
	Number Worse by 2 status zones	0	0	1	1	2
	Number Better by 1 status zone	0	0	13	7	2
	Number Better by 2 status zones	0	0	1	0	0

Table 12: **Contingency table of error types (None, Predicted Better, Predicted Worse) and confidence ratings (*Low, Medium, or High)* for WSP rapid statuses generated by the *Learning Tree 3* algorithm across all three status scales (see Table 4).** These are statuses assigned for the *learning data set* of 65 cases, which includes two assessments for Fraser sockeye CUs. The least precautionary outcome occurs where the WSP rapid status assigned by the algorithm is better than the WSP integrated status assessments *and* the confidence rating is *High*: this is highlighted in orange. Specifics for the fives cases where this least precautionary outcome occurred are summarized in Section 3.3. None: *Learning Tree 3* assigned an identical status to the WSP integrated status assigned for the same CU & data during expert workshops*;* Predicted Better: *Learning Tree 3* assigned a better status than the WSP integrated status*; Predicted Worse: Learning Tree 3* assigned a poorer status than WSP integrates status.

	Confidence Rating			
Error Type	**Low**	**Medium**	**High**	*Total*
None	3	26	25	**54**
Predicted Better	2	0	5	**7**
Predicted Worse	1	2	1	**4**
Total	**6**	**28**	**31**	**65**

FIGURES

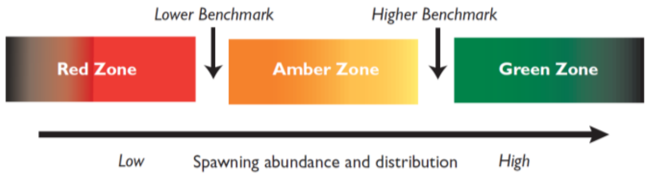


Figure 1. **Wild Salmon Policy status zones (*Red, Amber*, and *Green*) for individual status metrics.** Status zones delineated by lower and upper benchmarks from *Red*, which designates poor status, up to *Green*, which is a healthy status. Reprinted from Fisheries and Oceans Canada (2005). Note that statuses for individual metrics are combined into a single WSP integrated status assessment, which add two intermediate status zones (*Red/Amber, Amber/Green*; Table 1).

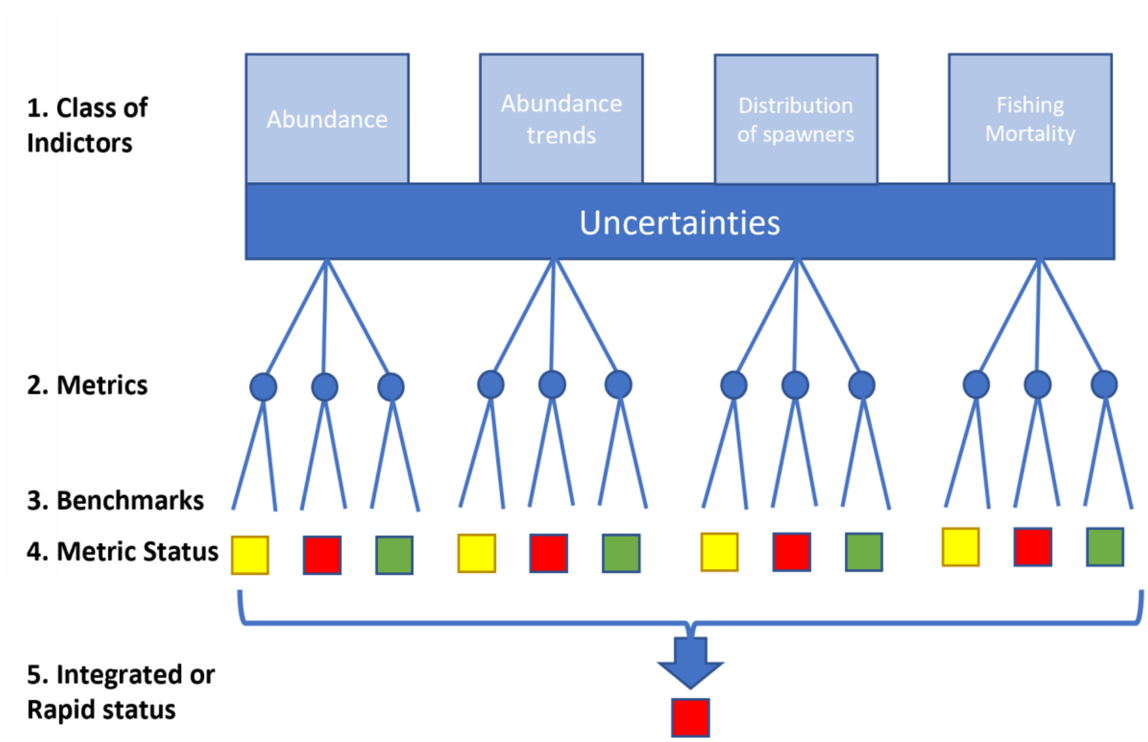


Figure 2. **Hierarchy for the assessment of biological status of CUs under the WSP.** Components include 1) four classes of indicators, 2) metrics within each indicator class, and 3) benchmarks on each metric; 4) statuses for each metric; 5) integrated or rapid statuses that rolls all the metric statuses into a single status for a CU. Revised from Holt et al. (2009).

R	A	G	Conservation Unit	Cyclic	Provisional
6			Takla-Trembleur – Estu	Y	1 Red
5			Nadina-Francois-ES		2 Red
5			Taseko-ES		
5			Nahatlatch-ES		
4			Bowron-ES		
5			Cultus-L		
6			Widgeon River – River Type		
2	4		Chilliwack-ES		1 Amber
2	4		Francois-Fraser-S		1 Red
5	1		Quesnel – S	Y	2 Red
2	2		Takla-Trembleur-Stuart-S	Y	
1	5		North Barriere-ES		1 Red
1	5		Anderson-Seton – ES	Y	
2	1	2	Seton-L (de novo)	Y	1 Amber
	5		Kamloops-ES		2 Amber
	5		Harrison (U/S)-L		1 Amber
	2	2	Pitt – ES		
	1	2	Shuswap – ES	Y	
		6	Chilko-S & Chilko-ES aggregate		2 Green
		4	Lillooet-Harrison-L		1 Green
		5	Harrison (D/S)-L		
1		4	Shuswap Complex – L	Y	
		5	Harrison River – River Type		

Figure 3. **Summary of group results for WSP integrated status assessments in the first Fraser sockeye WSP assessment.** This figure reproduces Table 3 from Grant and Pestal (2012). Participants worked in six groups. Status designations were labelled provisional if a group did not reach consensus. The majority view is shown in the columns on the left, but the number of groups with provisional status designations is also included. For example, '2 *Red*' for Quesnel-S means that two of the five groups that settled on a *Red* status had some dissenting views. By comparison, the '1' in the *Amber* column for Quesnel-S means that there was one group that reached a consensus designation of *Amber*, which could not be reconciled with the results from the other five groups through plenary discussion.

Figure 4. **Example of group results for WSP integrated status assessments in the Southern BC Chinook WSP assessment.** Participants worked in 6 groups, reviewing data summaries for a set of unidentified CUs, then posting the group results on the wall for a facilitated plenary discussion to determine a consensus status designation where possible. CU names were then revealed, and status designations finalized through further facilitated plenary discussions. In this example, the status assignments by individuals group mostly matched, and for all four cases there was a clear majority agreement on CU status, even though this set of CUs specifically included cases where trend metrics showed mixed signals (i.e. *long-term trend* and *percent change* metrics indicated different statuses).

SK 2010	Minimalist	FancyPants	CatReal	SimplyRed	LTree1	LTree2	LTree3	*IntScore5*	IntScore3	IntScore2	
ES_AndSet(CU-19)	A				G	A	A	A	A	NR	SK
ES_Bowr(CU-1)	R	R	R	R	R	R	R	R	R	R	SK
ES_Chill(CU-17)			R		R	R	R	RA	R	R	SK
ES_Kaml(CU-2)	A	A	A	NR	A	A	A	A	A	NR	SK
ES_NadFran(CU-3)	A	R	R	R	R	R	R	R	R	R	SK
ES_Naha(CU-14)	R	R		R	R	R	R	R	R	R	SK
ES_Nbarr(CU-4)	A	A	A	NR	A	A	A	A	A	NR	SK
ES_Pitt(CU-5)	A	A	A	NR	A	G	A	AG	A	NR	SK
ES_Shu(CU-20)	A				G	A	A	AG	A	NR	SK
ES_Tas(CU-15)	R	R	R	R	R	R	R	R	R	R	SK
Estu(CU-18)	R	R		R	R	R	R	R	R	R	SK
L_Cult(CU-8)	R	R	R	R	R	R	R	R	R	R	SK
L_HarrDS(CU-13)	G		A		G	G	G	G	G	NR	SK
L_HarrUS(CU-11)	R	R	A	R	A	A	A	A	A	NR	SK
L_LilHarr(CU-12)	A	G	A	NR	A	A	A	G	G	NR	SK
L_Shu(CU-23)	R	R		R	R	G	R	G	G	NR	SK
RT_HarrR(CU-9)	G	G	A	NR	G	G	G	G	G	NR	SK
RT_Wid(CU-16)	R	R		R	R	R	R	R	R	R	SK
S_Chilk(CU-6)	A	RA	A	R	G	G	G	G	G	NR	SK
S_FranFras(CU-7)	A	RA	A	NR	A	A	A	RA	R	R	SK
S_Ques(CU-22)	R	RA		R	A	A	R	RA	R	R	SK
S_TakTremStu(CU-21)	R	RA		R	A	A	R	RA	R	R	SK

Figure 5. **WSP Rapid statuses for each Fraser sockeye CU (rows) and candidate algorithms (first seven columns), compared to WSP integrated statuses using data up to 2010 (last 3 columns) (Grant & Pestal 2012).** The WSP integrated statuses were assigned on a 5 status scale (marked with *) (Grant and Pestal 2012); these were also converted to a 3 and 2 status scales for better comparisons to particular algorithms (see Table 4 to Table 6 for details). For some cases, all or most algorithms could assign a status and the rapid statuses match the WSP integrated assessment (e.g. Bowron_ES, Cultus_L, Nahatlatch-ES, Takla_Trembleur_EStu,Taseko_ES; Widgeon_RT). For others, different algorithms and WSP integrated statuses spanned multiple status zones (e.g. Harrison_US_L).

SK 2015	Minimalist	FancyPants	CatReal	SimplyRed	LTree1	LTree2	LTree3	*IntScore5*	IntScore3	IntScore2	
ES_AndSet(CU-19)	G	AG	A	NR	A	A	A	AG	A	NR	SK
ES_Bowr(CU-1)	R	R	R	R	R	R	R	R	R	R	SK
ES_Chill(CU-17)	A	AG	A	NR	A	A	A	AG	A	NR	SK
ES_Kaml(CU-2)	A	A	A	NR	A	A	A	A	A	NR	SK
ES_NadFran(CU-3)	A	AG	A	NR	A	A	A	AG	A	NR	SK
ES_Naha(CU-14)	A				G	G	A	A	A	NR	SK
ES_Nbarr(CU-4)	A	A	A	NR	A	A	A	A	A	NR	SK
ES_Pitt(CU-5)	G	G	A	NR	G	G	G	G	G	NR	SK
ES_Shu(CU-20)	A	AG	A	NR	A	A	A	A	A	NR	SK
ES_Tas(CU-15)	R	R	R	R	R	R	R	R	R	R	SK
Estu(CU-18)	R	R	R	R	R	R	R	R	R	R	SK
L_Cult(CU-8)	R	R	R	R	R	R	R	R	R	R	SK
L_HarrDS(CU-13)	G		A	R	A	A	R	AG	A	NR	SK
L_HarrUS(CU-11)	R	R	R	R	R	R	R	R	R	R	SK
L_LilHarr(CU-12)	A	A	A	R	A	A	A	A	A	NR	SK
L_Set(CU-10)	R	R	R	R	R	R	R	R	R	R	SK
L_Shu(CU-23)	R	R	A	R	G	G	G	AG	A	NR	SK
RT_HarrR(CU-9)	G	G	A	NR	G	G	G	G	G	NR	SK
RT_Wid(CU-16)	A				R	R	R	R	R	R	SK
S_Chilk(CU-6)	G	G	A	NR	G	G	G	G	G	NR	SK
S_FranFras(CU-7)	A	G	A	NR	G	G	G	AG	A	NR	SK
S_Ques(CU-22)	R	RA	R	R	R	R	R	RA	R	R	SK
S_TakTremStu(CU-21)	A	RA	A	NR	A	A	A	RA	R	R	SK

Figure 6. **WSP rapid statuses for each Fraser sockeye CU (rows) and candidate algorithms (first seven columns), compared to WSP integrated statuses using data up to 2015 (last 3 columns) (Grant et al. 2020).** The WSP integrated statuses were assigned on a 5 status scale (marked with *) (Grant et al. 2020); these were also converted to a 3 and 2 status scales for better comparisons to particular algorithms (see Table 4 to Table 6 for details). For some cases, all or most algorithms could assign a status and the rapid statuses match the WSP integrated assessment (e.g. Bowron_ES, Cultus_L,). For others, different algorithms and WSP integrated statuses spanned multiple status zones (e.g. Shuswap_L).

CK 2012 / CO 2013

	Minimalist	FancyPants	CatReal	SimplyRed	LTree1	LTree2	LTree3	IntScore5	*IntScore3*	IntScore2	
LFR-fall(CK-03)	A	G	A	NR	G	G	G	G	G	NR	CK
LTh(CK-17)	R	R	R	R	R	R	R	R	R	R	CK
MFR-spring(CK-10)	R	R	R	R	R	R	R	R	R	R	CK
MFR-summer(CK-11)	A		A		G	A	A	A	A	NR	CK
NEVI(CK-29)	A		A		G	G	A	R	R	R	CK
NoKy(CK-32)	R	R	R	R	R	R	R	R	R	R	CK
NTh-spr(CK-18)	R	R	R	R	R	R	R	R	R	R	CK
NTh-sum(CK-19)	R	R	R	R	R	R	R	R	R	R	CK
OK(CK-01)	R	R	R	R	R	R	R	R	R	R	CK
Portage(CK-09)	R	R	A	R	R	A	R	R	R	R	CK
STh-0.3(CK-13)	A		A		G	G	A	G	G	NR	CK
STh-1.3(CK-14)	A		A		G	A	A	RA	R	R	CK
STh-BESS(CK-16)	R	R	R	R	R	R	R	R	R	R	CK
SWVI(CK-31)	R	R	R	R	R	R	R	R	R	R	CK
UFR-spring(CK-12)	R	R	R	R	R	R	R	R	R	R	CK
FrCny(CO-02)	A	A	A	NR	G	G	A	A	A	NR	CO
Lthom(CO-03)	A	AG	A	NR	G	G	A	AG	A	NR	CO
MFr(CO-01)	A	A	A	NR	G	G	A	A	A	NR	CO
NTh(CO-04)	A	AG	A	NR	G	G	G	AG	A	NR	CO
STh(CO-05)	A	A	A	NR	G	G	A	A	A	NR	CO

Figure 7. **WSP rapid statuses for each Southern BC Chinook using data up to 2012 (DFO 2016) and Interior Fraser coho CUs (using data up to 2013 (DFO 2015) (rows) and candidate algorithms (first seven columns), compared to WSP integrated statuses (last 3 columns)**. The WSP integrated statuses were assigned on a 5 status scale (marked with *); these were also converted to a 3 and 2 status scales for better comparisons to particular algorithms (see Table 4 to Table 6 for details).

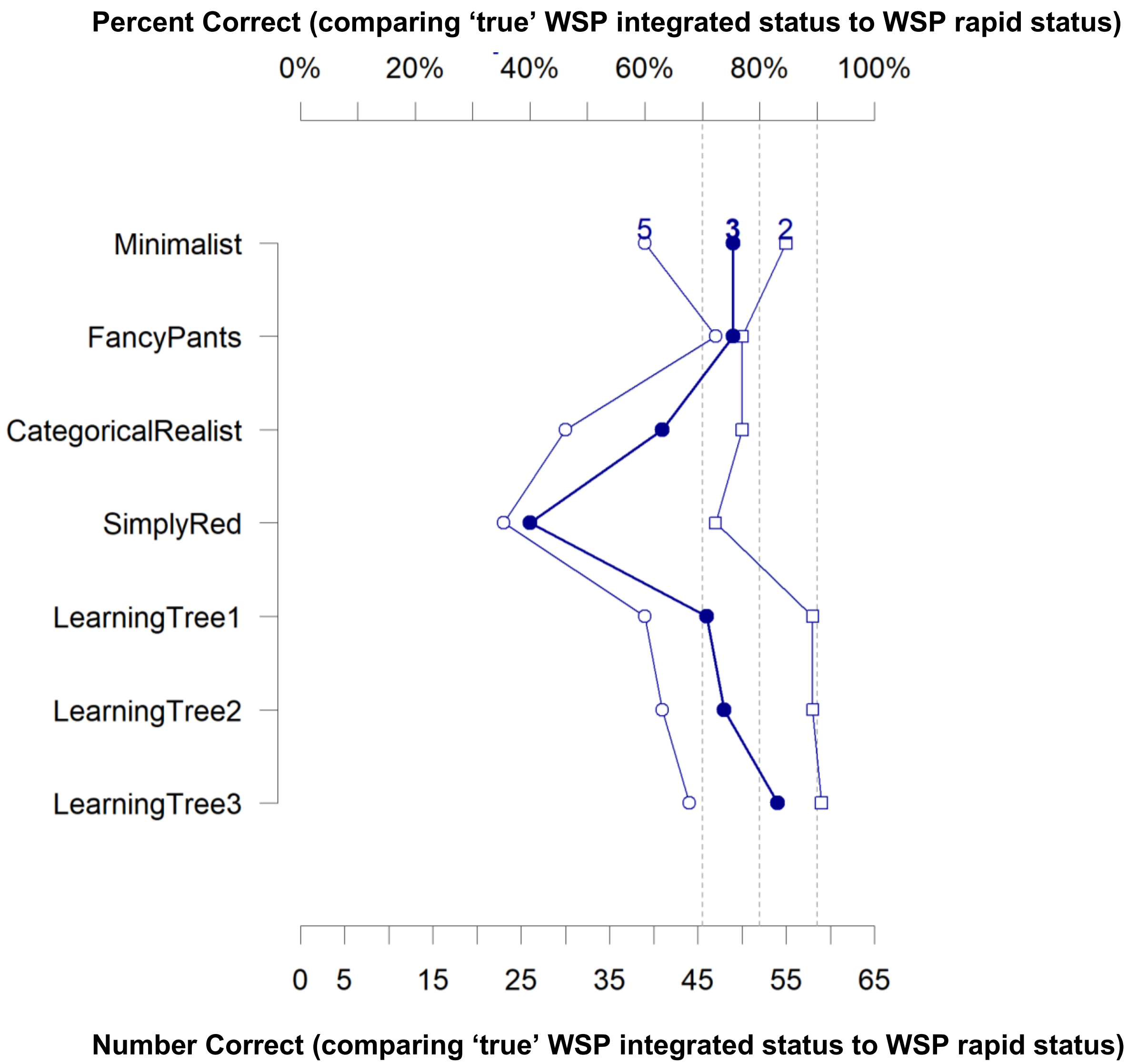


Figure 8. **Algorithm comparison based on correct rapid status designations.** Number (lower horizontal axis) or percent (upper horizontal axis) of correct WSP rapid statuses, as compared to WSP integrated statuses, out of the total 65 cases in the *learning data set* for three fitted algorithms (*Minimalist, Fancy Pants & Categorical Realist;* and four constructed algorithms: *Simply Red, LearningTree1, LearningTree2, LearningTree3)*. Results are shown for the three alternative status scales (5,3, and 2), as explained Table 4. This is one out of several performance measures used; the full set are presented in Table 7. Candidate algorithms were evaluated against criteria with a combination of quantitative and qualitative performance measures, not exclusively based on this figure.

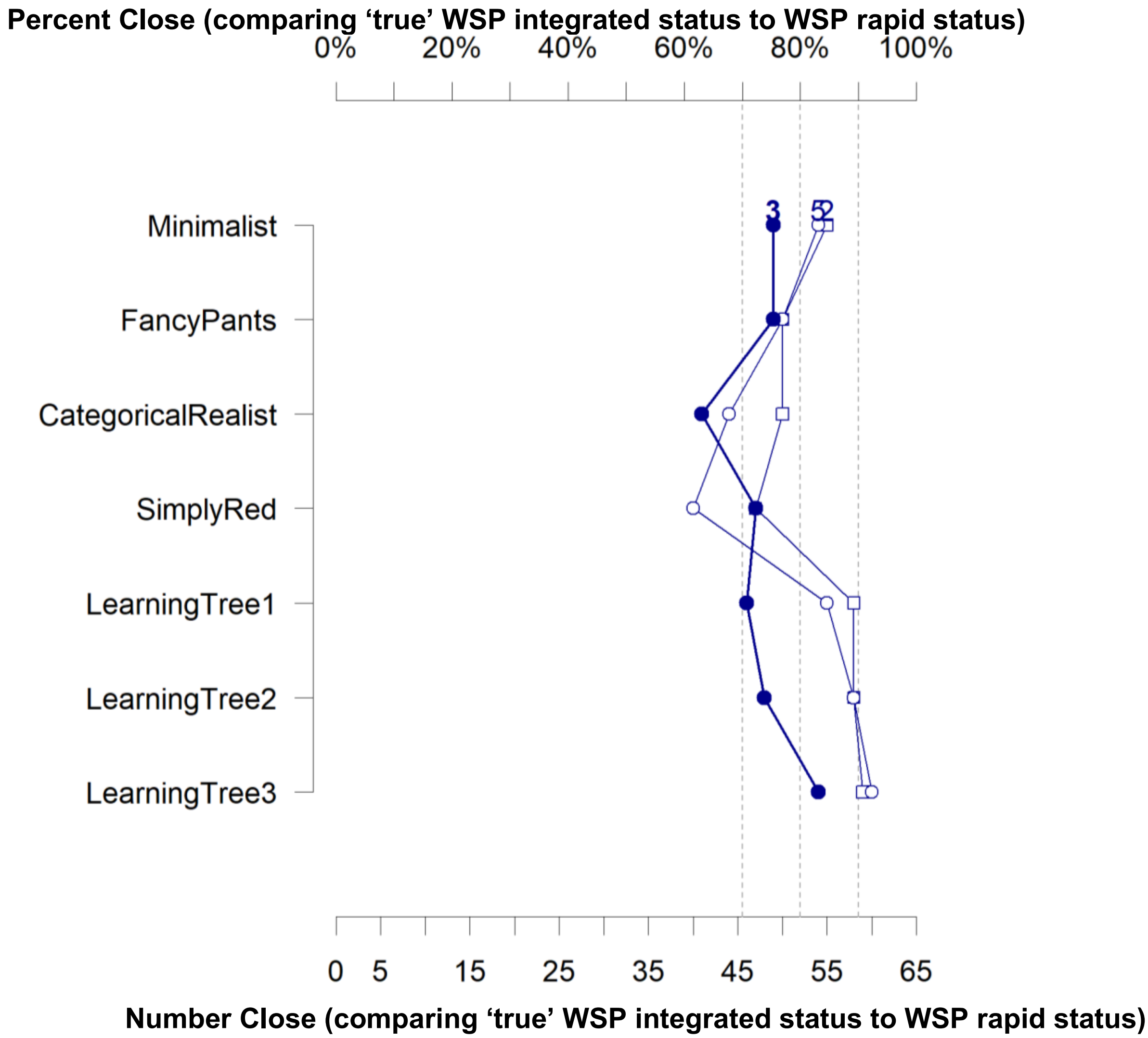


Figure 9. **Algorithm comparison based on close rapid status designations**. Number (lower horizontal axis) or percent (upper horizontal axis) of close WSP rapid statuses, as compared to WSP integrated statuses, out of the total 65 cases in the learning data set for three fitted algorithms (Minimalist, Fancy Pants & Categorical Realist; and four constructed algorithms: Simply Red, LearningTree1, LearningTree2, LearningTree3. Close indicates the WSP rapid status in only 1 status zone different from the WSP integrated status on a 5 status scale. For example, a CU assessed as Amber by the expert process and assigned Not Red by the Simply Red algorithm would be scored as incorrect in Figure 2, but scored as “close” in this figure. Results are shown for the three alternative WSP status scales (5,3, or 2 status categories), as explained in Table 4. This is one out of several performance measures used; the full set are presented in Table 7. Candidate algorithms were evaluated against criteria with a combination of quantitative and qualitative performance measures, not exclusively based on this figure.

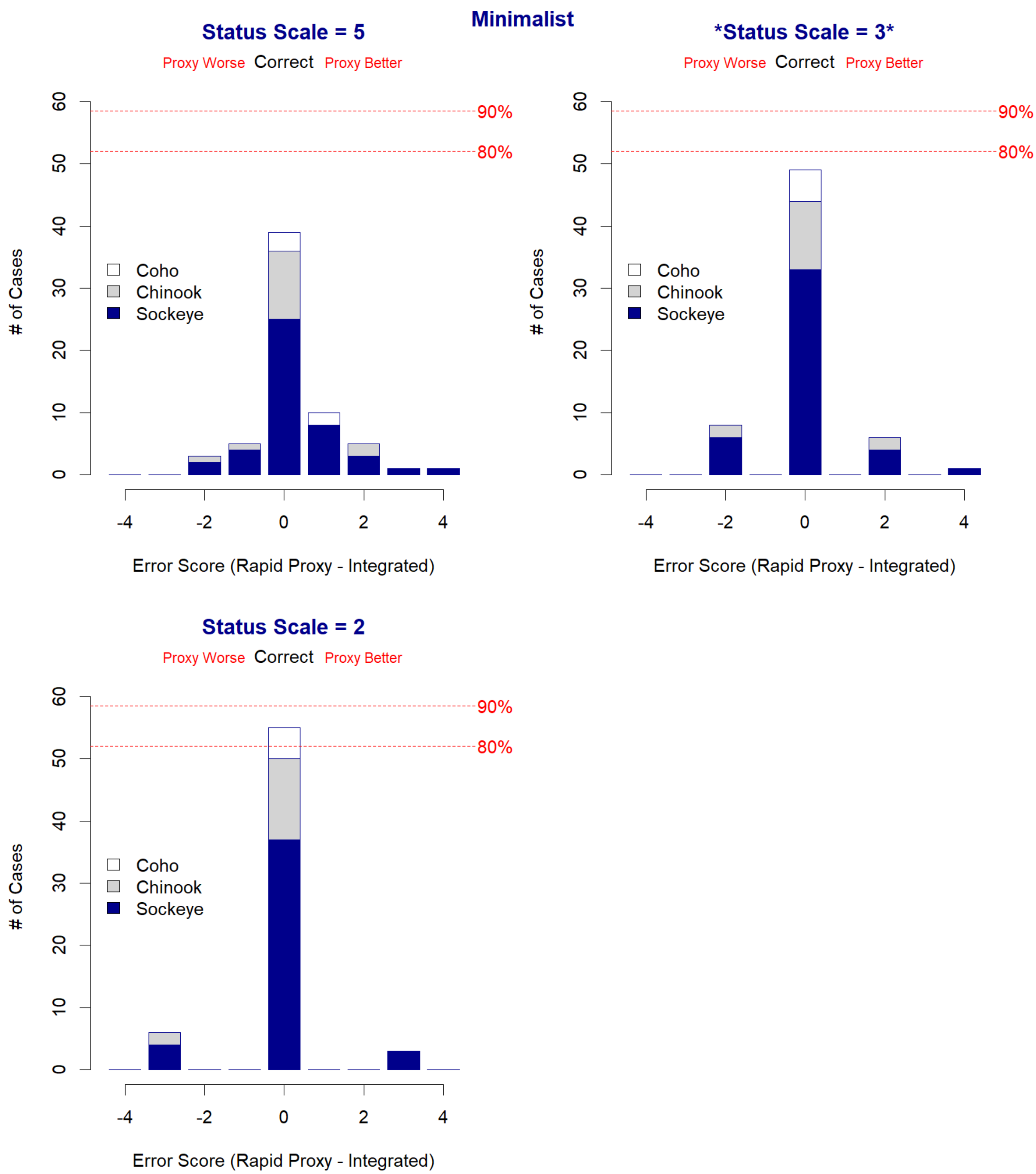


Figure 10. ***Minimalist* algorithm: distribution of errors in *learning data set* statuses.** Each panel shows the frequency distribution of errors for the *completed* cases (i.e. number of cases for each type of error from the cases for which a WSP rapid status could be assigned with that algorithm). Errors were calculated on a 5,3, or 2 status scale, as described in Table 4 to Table 6.
Table 7 to Table 10 list the corresponding values. The error scale for which the algorithm was designed is marked by * in the panel title.

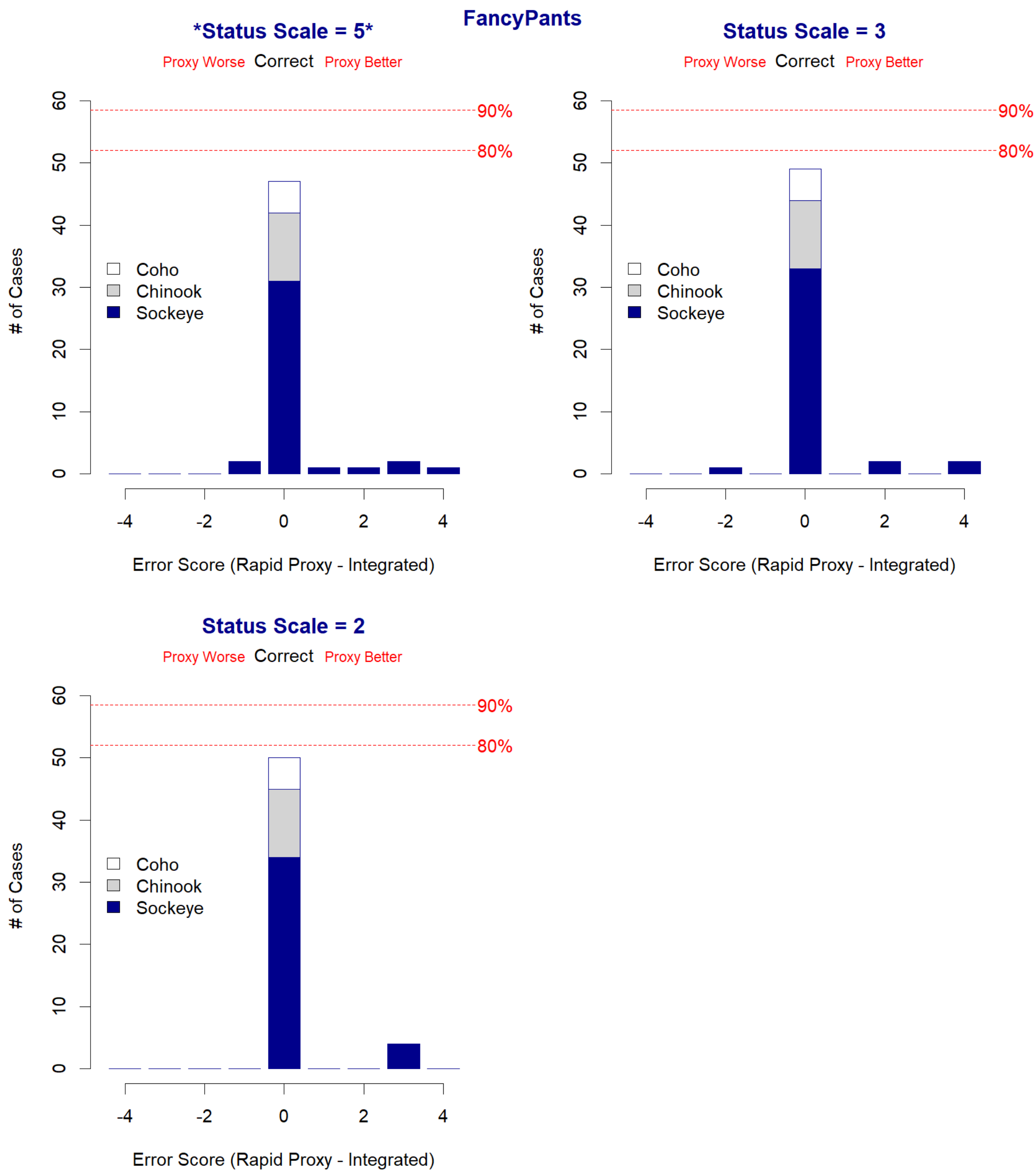


Figure 11. ***Fancy Pants* algorithm: distribution of errors in *learning data set* status approximations.** Each panel shows the frequency distribution of errors for the *completed* cases (i.e. number of cases for each type of error from the cases for which a WSP rapid status could be assigned with that algorithm). Errors were calculated on a 5,3, or 2 status scale, as described in Table 4 to Table 6.
Table 7 to Table 10 list the corresponding values. The error scale for which the algorithm was designed is marked by * in the panel title.

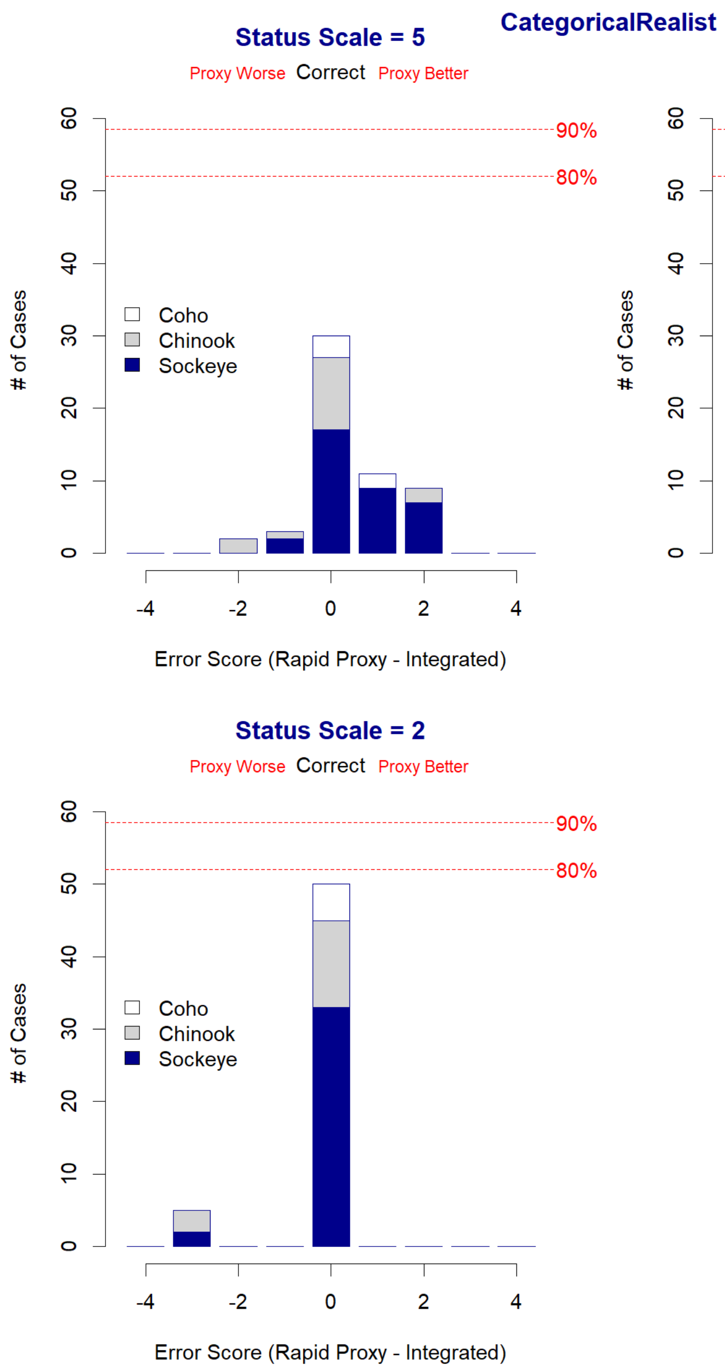


Figure 12. ***Categorical realist* algorithm: distribution of errors in *learning data set* status.** Each panel shows the frequency distribution of errors for the *completed* cases (i.e. number of cases for each type of error from the cases for which a WSP rapid status could be assigned with that algorithm). Errors were calculated on a 5,3, or 2 status scale, as described in Table 4 to Table 6.
Table 7 to Table 10 list the corresponding values. The error scale for which the algorithm was designed is marked by * in the panel title.

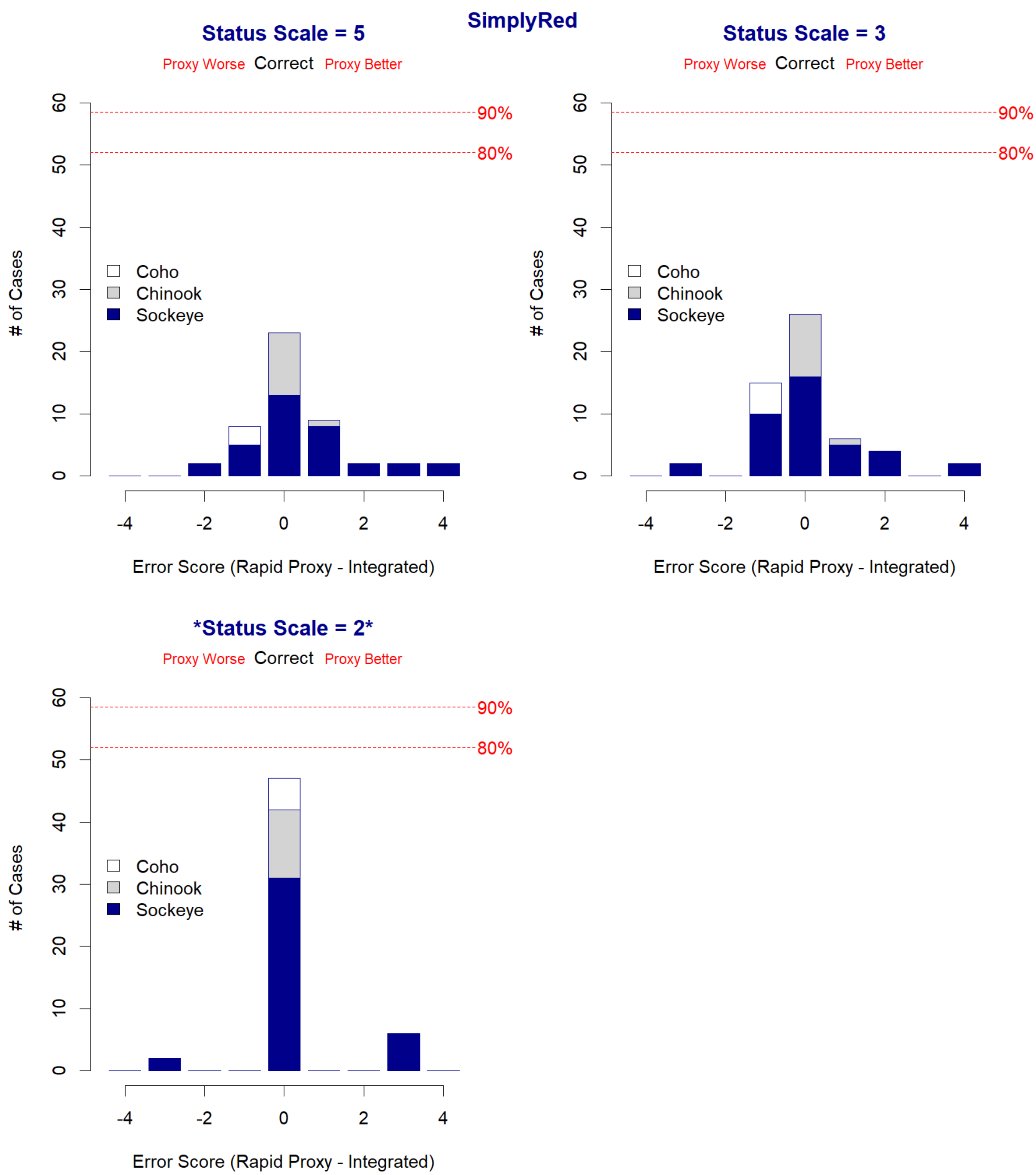


Figure 13. ***Simply Red* algorithm: distribution of errors in learning data set status.** Each panel shows the frequency distribution of errors for the completed cases (i.e. number of cases for each type of error from the cases for which a WSP rapid status could be assigned with that algorithm). Errors were calculated on a 5,3, or 2 status scale, as described in Table 4 to Table 6.
Table 7 to Table 10 list the corresponding values. The error scale for which the algorithm was designed is marked by * in the panel title.

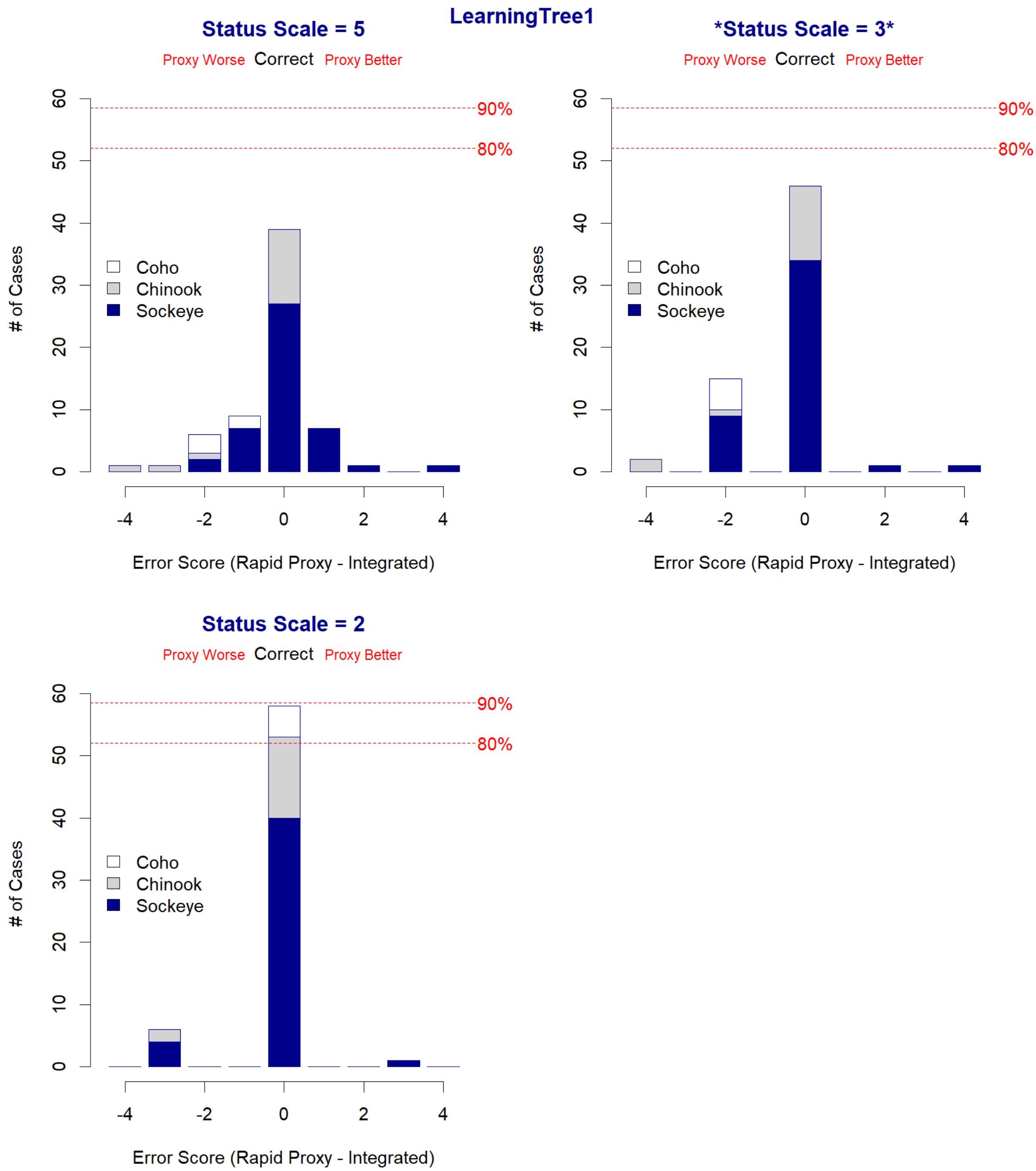


Figure 14. ***Learning Tree 1* algorithm: distribution of errors in *learning data set* status.** Each panel shows the frequency distribution of errors for the *completed* cases (i.e. number of cases for each type of error from the cases for which a WSP rapid status could be assigned with that algorithm). Errors were calculated on a 5,3, or 2 status scale, as described in Table 4 to Table 6.
Table 7 to Table 10 list the corresponding values. The error scale for which the algorithm was designed is marked by * in the panel title.

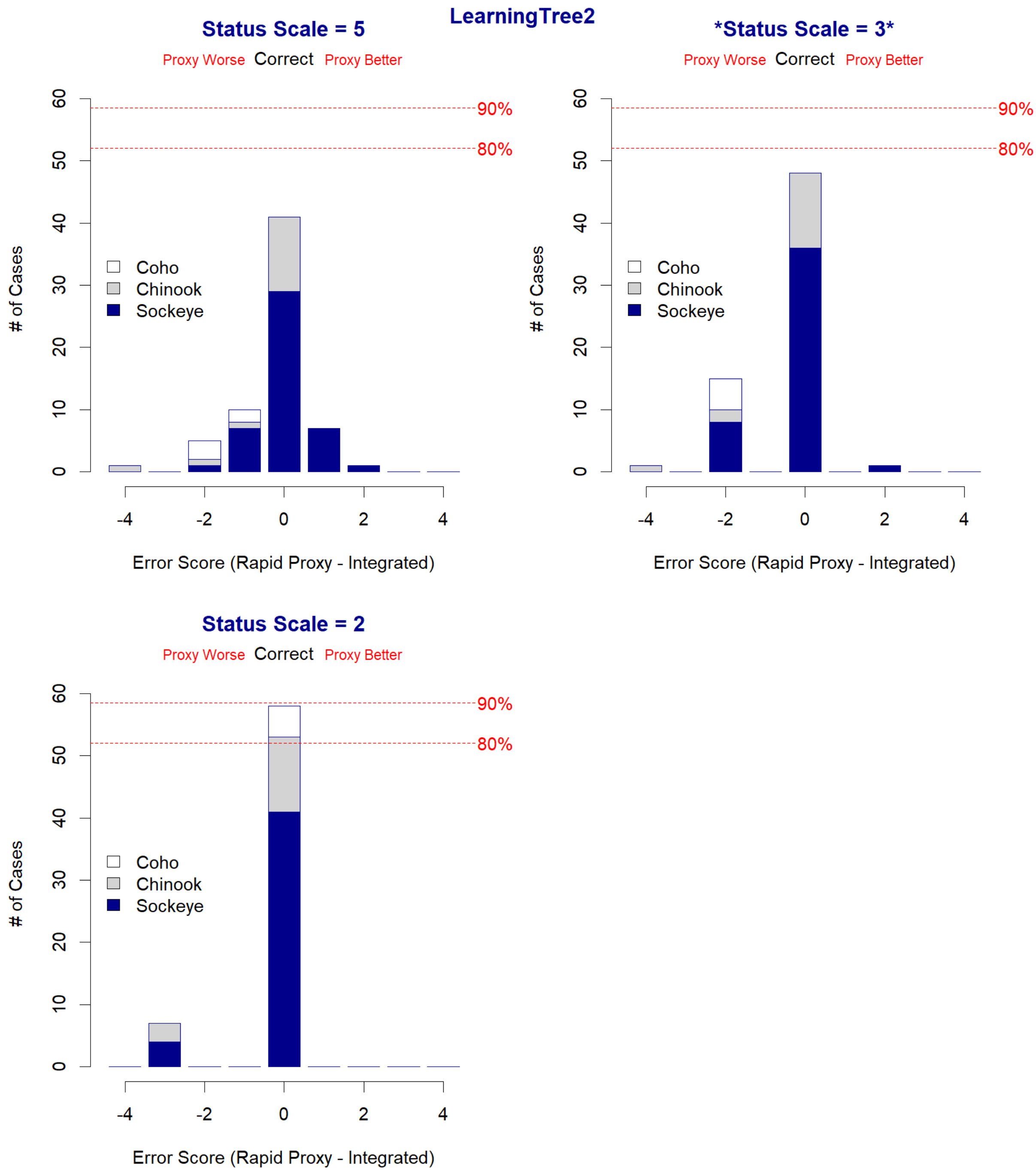


Figure 15. ***Learning Tree 2* algorithm: distribution of errors in *learning data set* status.** Each panel shows the frequency distribution of errors for the *completed* cases (i.e. number of cases for each type of error from the cases for which a WSP rapid status could be assigned with that algorithm). Errors were calculated on a 5,3, or 2 status scale, as described in Table 4 to Table 6.
Table 7 to Table 10 list the corresponding values. The error scale for which the algorithm was

designed is marked by * in the panel title.

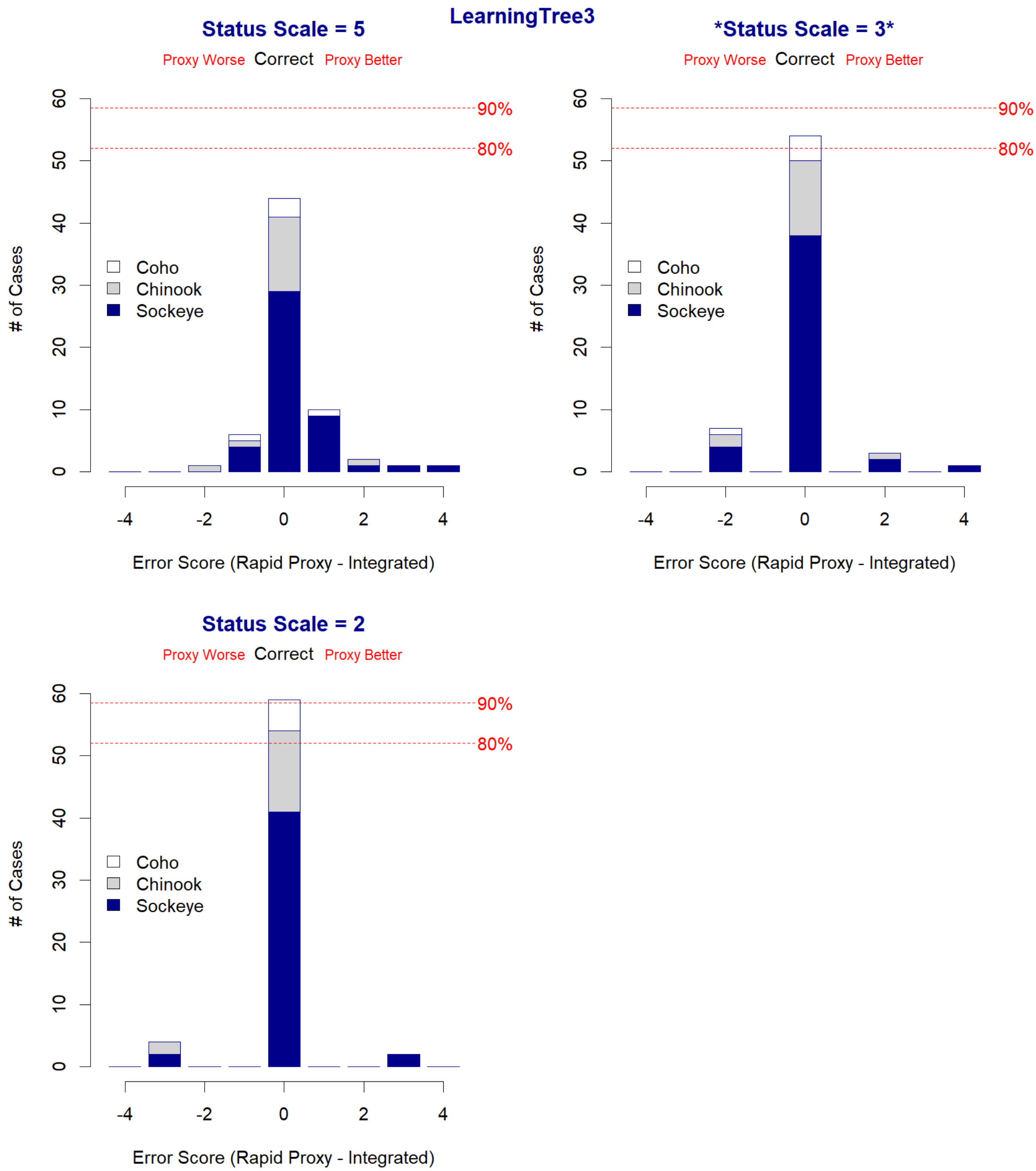


Figure 16. ***Learning Tree 3* algorithm: distribution of errors in *learning data set* status.** Each panel shows the frequency distribution of errors for the *completed* cases (i.e. number of cases for each type of error from the cases for which a WSP rapid status could be assigned with that algorithm). Errors were calculated on a 5,3, or 2 status scale, as described in Table 4 to Table 6.
Table 7 to Table 10 list the corresponding values. The error scale for which the algorithm was designed is marked by * in the panel title.

	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019
Widgeon_RT	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R
Taseko_ES	N	N	R	R	R	R	R				A	R	R	R	R	R	R	R	R	R	R	R	R	R	N
Takla_Trem_S_S	G	G	G	G	G	G	A	A	A	A	A	A	A	A	R	R	R	R	A	A	A	A	A	A	A
Takla_Trem_EStu	G	G	A	A	A	A	A	A	A	A	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R
Shuswap_L	A	A	A	A	A	A	A	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	A	A	A
Shuswap_ES	A	A	A	A	A	A	A	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	A
Seton_L	A	A	A	A	A	A	A	A	A	A	A	A	A	A	R	R	R	R	R	R	R	R	R	R	R
Quesnel_S	G	G	G	G	G	G	G	G	G	G	A	A	A	A	R	R	R	R	R	R	R	R	R	R	R
Pitt_ES	A	A	A	G	G	G	G	G	G	G	G	G	G	G	G	A	G	G	G	G	G	G	G	G	A
North_Barriere_ES	A	G	G	G	G	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	R	R
Nahatlatch_ES	A	A	A	A	A	A	A	A	A	A	A	A	A	R	R	R	R	R	A	A	A	A	A	A	R
Nadina_Francois_ES	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	A	A	R	A	A
Lillooet_Harr_L	G	A	A	A	A	A	A	A	A	G	G	G	A	A	A	A	A	A	A	A	A	A	A	A	R
Kamloops_ES	R	R	R	A	A	A	A	G	G	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	R
Harrison_US_L	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	R	R	R	R	R
Harrison_R	R	R	R	R	R	R	R	R	R	R	R	R	A	A	A	A	G	G	G	G	G	G	A	A	R
Harrison_DS_L	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	R	R	G	G	G	G
Fran_Fras_S	A	A	A	A	G	G	G	G	G	A	G	A	A	A	A	A	A	A	A	A	G	A	A	A	A
Chilliwack_ES	N	N	N	N	N	N	N	N	N	A	R	R	R	R	R	R	R	A	A	A	A	A	A	R	R
Chilko_S_ES	G	G	G	G	G	G	G	G	G	A	A	A	A	A	A	A	G	G	G	G	G	G	G	A	A
Bowron_ES	R	A	A	A	A	A	A	A	A	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R
Anderson_Seton_ES	A	A	A	A	A	A	A	A	A	A	A	A	A	A	R	A	A	A	A	A	A	A	A	A	A

Figure 17. **Retrospective pattern of *Learning Tree 3* WSP rapid statuses (rows) for Fraser sockeye CUs (columns) from 1995 to 2019.** This figure illustrates the retrospective application of one algorithm to a subset of the CUs in the *learning data set*. It shows how statuses change several times since 1995 for some CUs (e.g. Harrison_R), while others are consistent for all or most of that time period (e.g. Widgeon_RT, Anderson_Seton_ES), and still others show a directional change to worsening status (e.g. Seton_L, Harrison_US_L, Takla-Trembleur_EStu). Appendices D and E show detailed retrospective results for all the CUs and candidate algorithms. Figure 18 and Figure 19 summarize the pattern over time across all CUs.

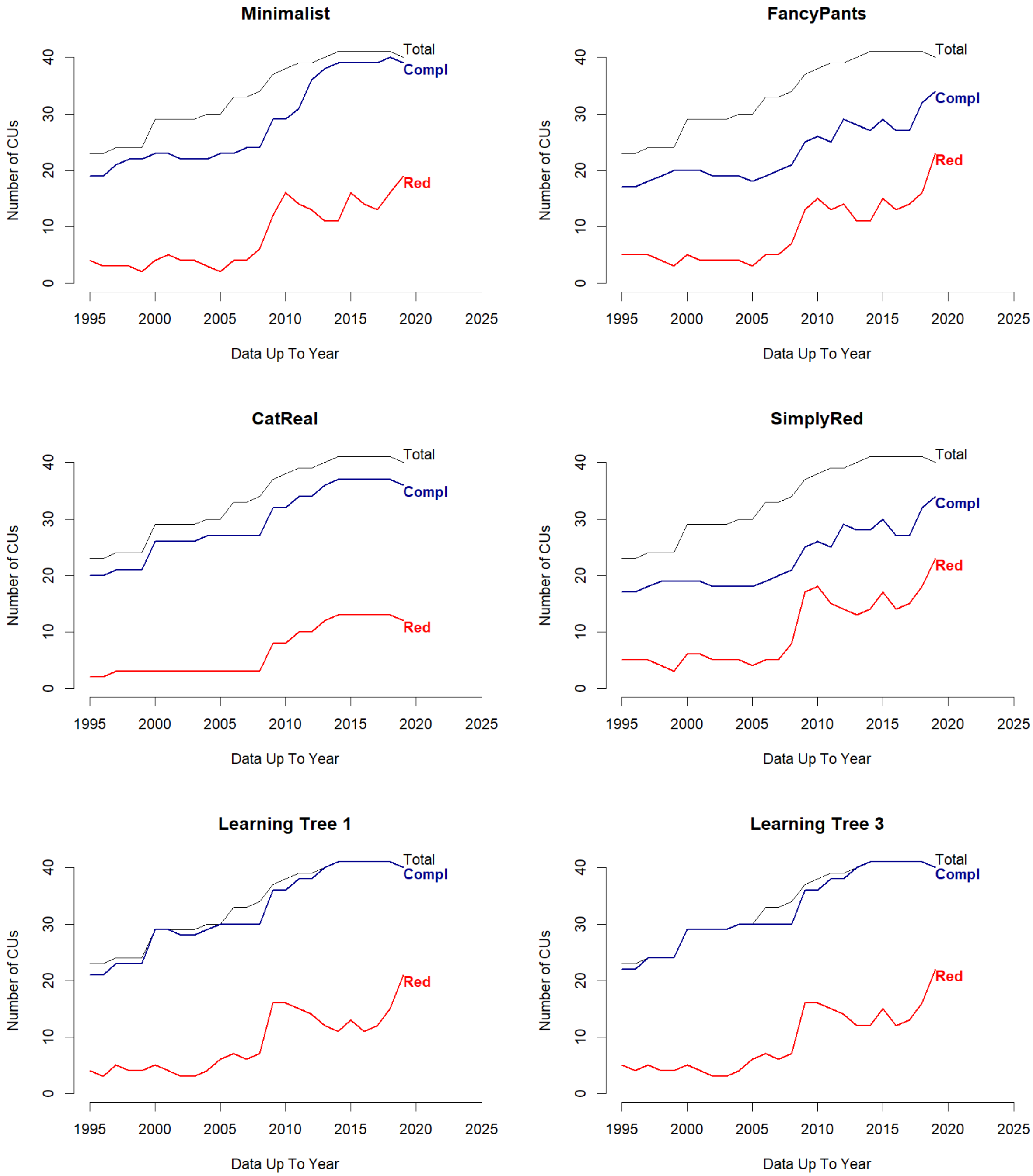


Figure 18. **Summary of retrospective WSP rapid statuses – number of completed CUs and number of *Red* statuses for Fraser sockeye, Southern BC Chinook and Interior Fraser coho CUs.** *Total* is the number of CUs with at least one status metric available for that year. *Compl* is the number of completed WSP rapid statuses assigned. *Red* is the number of CUs assigned a *Red* status by that algorithm. The retrospective test started in 1995. Earlier spawner data are available for many of the CUs, but metric availability varies a lot for the earlier parts of the time series due to individual metric requirements, for example, at least 3 generations of data (12-15 years) are needed to calculate the *percent change* metric.

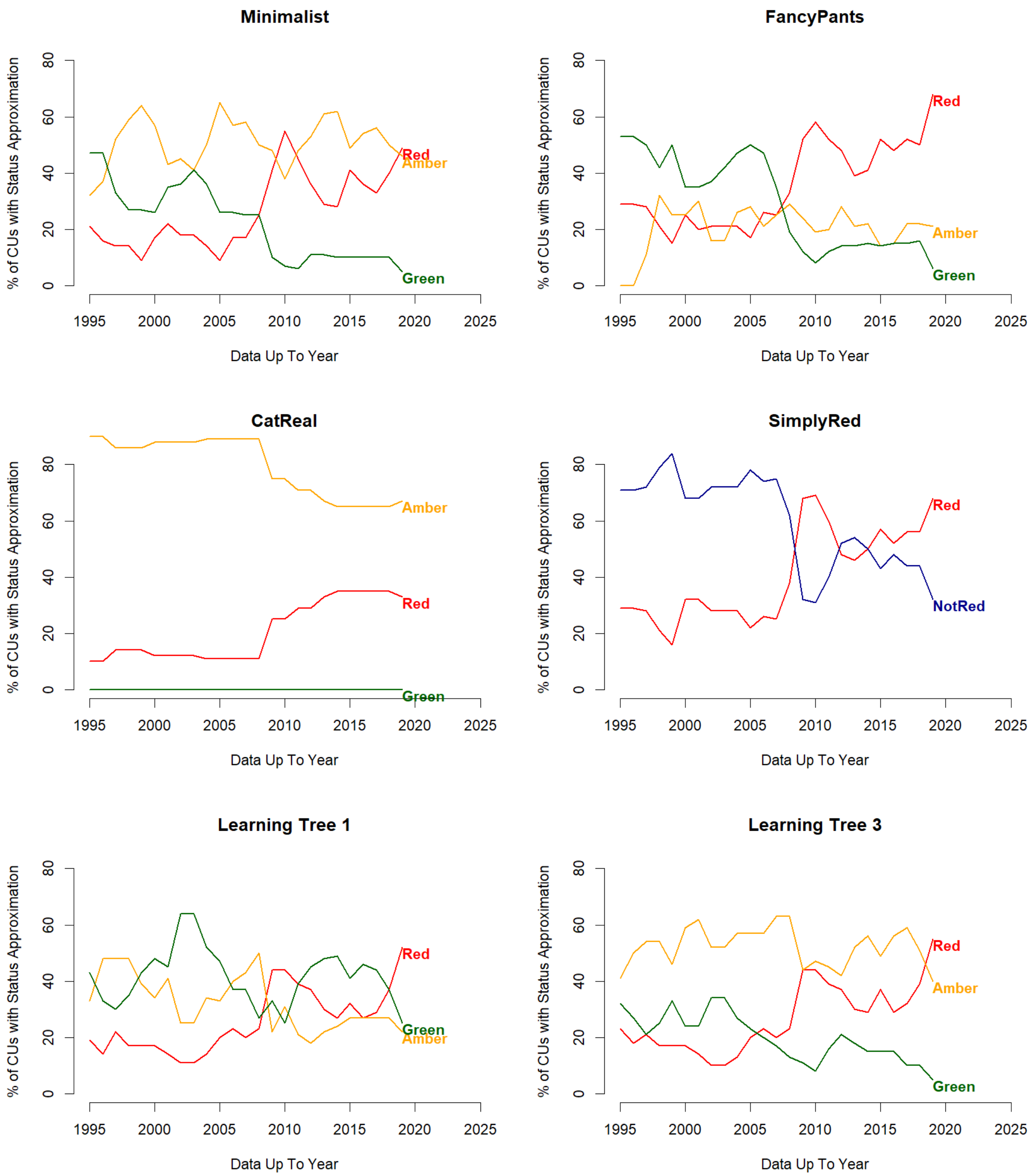


Figure 19. **Summary of retrospective WSP status – percent assigned to each status category for Fraser sockeye, Southern BC Chinook and Interior Fraser coho CUs.** Figure shows the percent of completed cases (i.e. number of cases assigned to each status category from the cases for which a WSP rapid status could be assigned with that algorithm). Refer to Figure 18 for the number and percent of cases that were completed. The proportion of CUs assigned to *Red* status follows the same pattern for all the algorithms except the *Categorical Realist*. *Amber* and *Amber* status assignments, however, differ between algorithms. The retrospective test started in 1995. Earlier spawner data are available for many CUs, but metric availability varies a lot for the earlier parts of the time series due to individual metric requirements, for example, at least 3 generations of data (12-15 years) are needed to calculate the *percent change* metric.

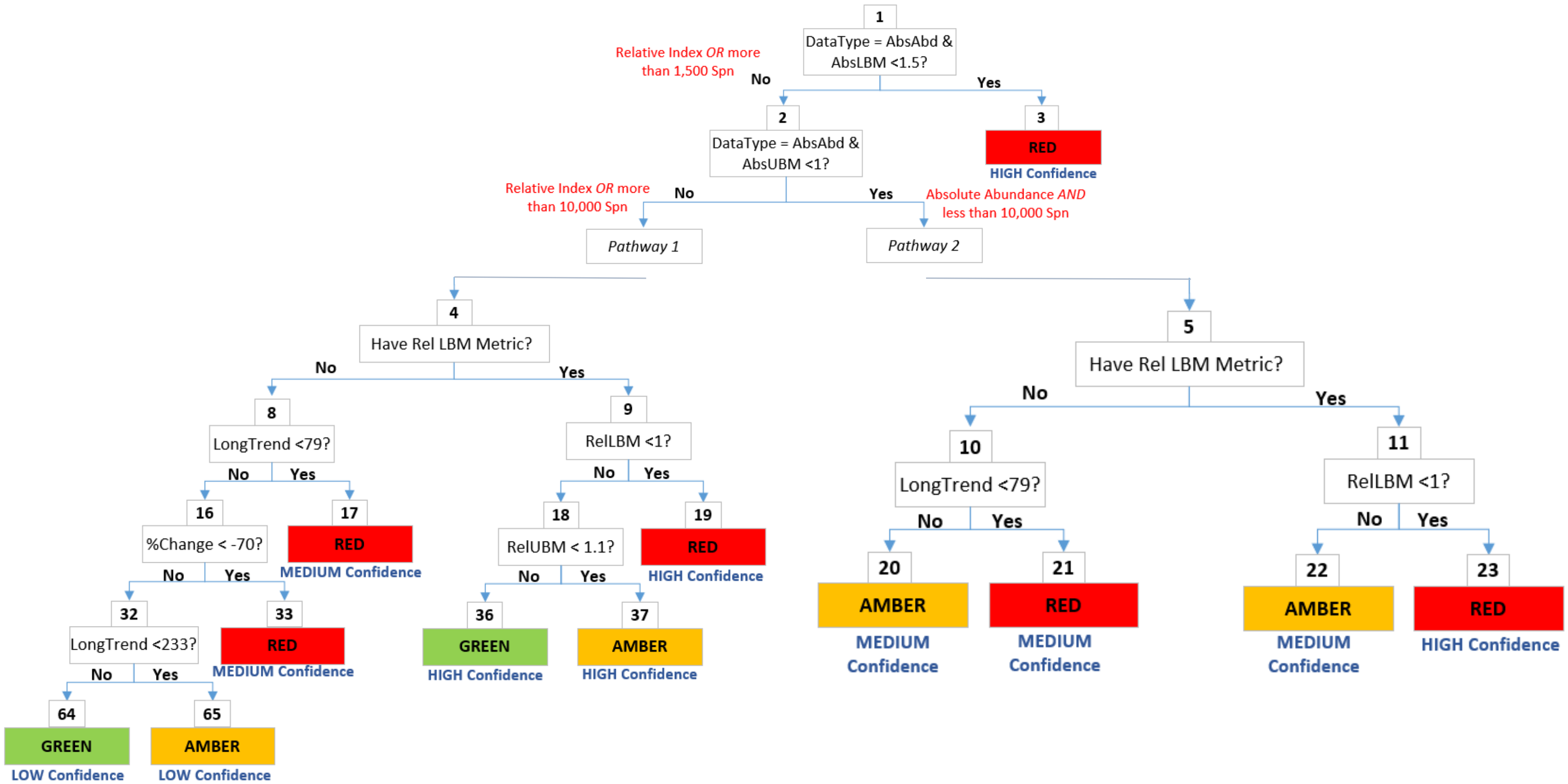


Figure 20. **WSP rapid status *Learning Tree 3* algorithm**. This is the selected rapid status algorithm, where a CU's metric values are compared to algorithm thresholds to determine final rapid status. Yes or no answers to these different decision points split the paths on the decision tree, terminating at rapid status assignments of *Red*, *Amber* or *Green*. The different splits in each pathway are identified as nodes, numbered from 1 to 65. **Pathway 1** is taken when the CU has no absolute abundance data, or these data exist, but fall above its upper benchmark of 10,000. **Pathway 2** is taken when the CU has absolute abundance data and these fall under its upper benchmark of 10,000. AbsAbd: absolute abundance; AbsLBM: *absolute abundance* lower benchmark; AbsUBM: *absolute abundance* upper benchmark; Rel LBM: *relative abundance* lower benchmark; LongTrend: is *long-term trend* metric; %Change: *percent change* metric; RelUBM: *relative abundance* upper benchmark. Confidence ratings for each end node are shown below the node. Section 2.5 describes the rationale for the confidence ratings.

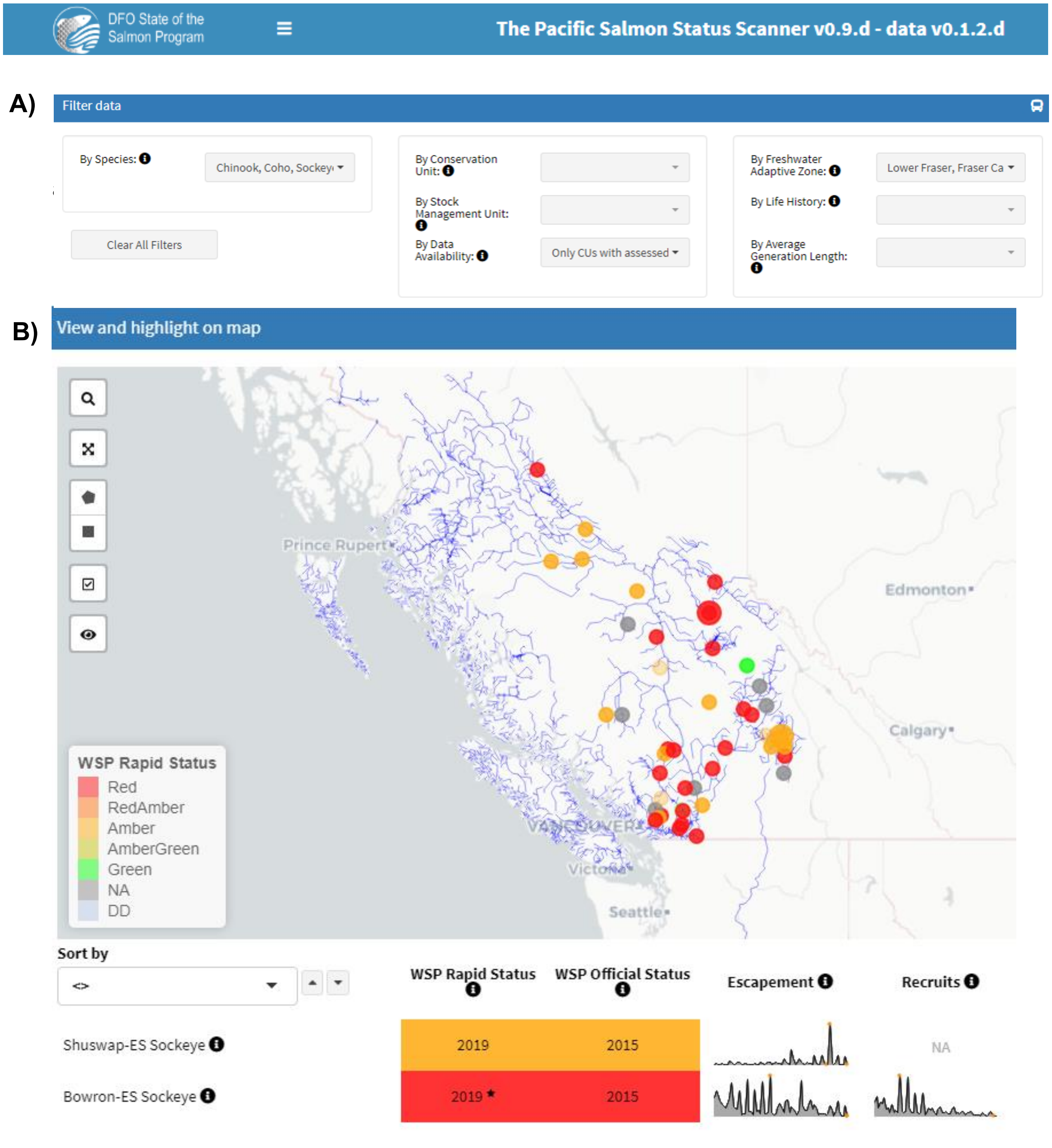
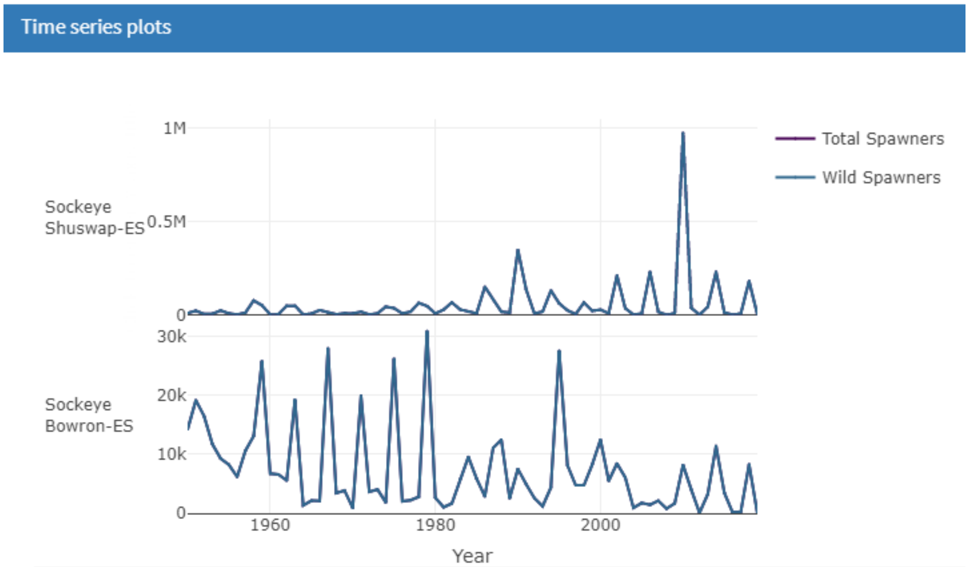


Figure 21. **Screen captures of DFO's Salmon Scanner in use. A)** In the filter data tab, users choose which data they want to work with by selecting attributes from seven drop-down lists. Here we have selected only CUs with assessed data from the Fraser watershed. **B)** In 'view and highlight on map' users can interact with filtered salmon CUs to view annual statuses and other attributes; they may also highlight CUs to explore more detailed information. We show CUs colour-coded by their 2019 WSP rapid statuses, and have highlighted two Fraser sockeye CUs, shown in the table below the map.

C)



D)

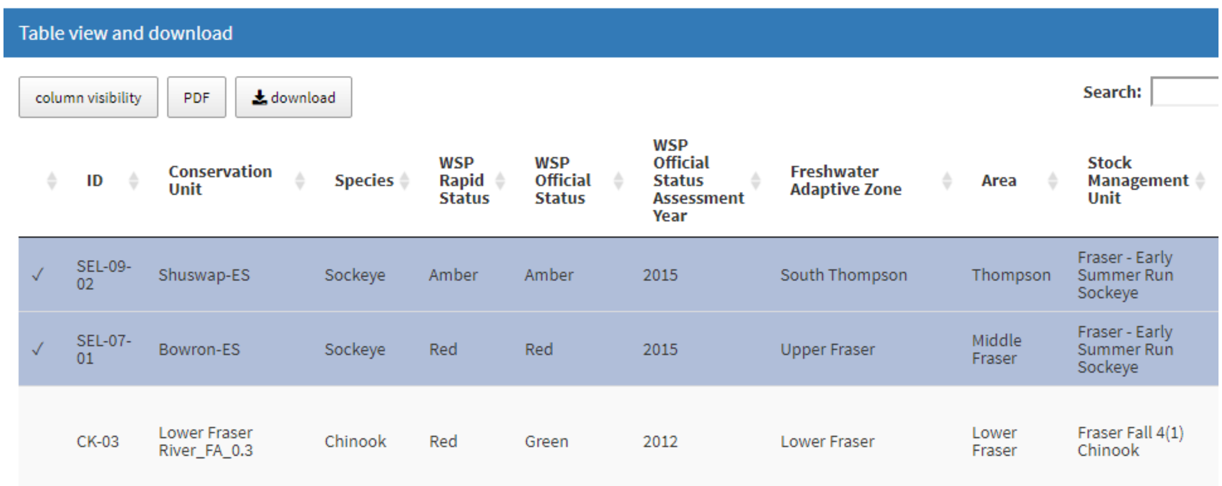


Table view and download

column visibility | PDF | download | Search:

	ID	Conservation Unit	Species	WSP Rapid Status	WSP Official Status	WSP Official Status Assessment Year	Freshwater Adaptive Zone	Area	Stock Management Unit
✓	SEL-09-02	Shuswap-ES	Sockeye	Amber	Amber	2015	South Thompson	Thompson	Fraser - Early Summer Run Sockeye
✓	SEL-07-01	Bowron-ES	Sockeye	Red	Red	2015	Upper Fraser	Middle Fraser	Fraser - Early Summer Run Sockeye
	CK-03	Lower Fraser River_FA_0.3	Chinook	Red	Green	2012	Lower Fraser	Lower Fraser	Fraser Fall 4(1) Chinook

Figure 22. **Screen captures of DFO's Pacific Salmon Status Scanner in use. C)** Time series plots are shown for the highlighted CUs. Users can select which time series to view, can interact with plots, and can alter plot attributes to create print-quality figures for download. **D)** Table view and download shows the highlighted CUs alongside all of the data filtered into use. Users can create pdfs of this table, and can download the attribute table or the underlying datasets as .csv files.

APPENDIX A: WSP STRATEGY 1: STANDARDIZED MONITORING OF WILD SALMON STATUS

A.1 BACKGROUND

The *first core principle* is that WSP CUs were identified and rapid statuses were developed based on conservation biology principles and are aligned with scientific peer-reviewed publications. This Appendix specifically covers the alignment of the WSP rapid status approach with the WSP and scientific peer-reviewed publications.

DFO's WSP recognizes the importance of salmon biodiversity (DFO 2005). Greater biodiversity reduces the risk of extinction by increasing the likelihood that salmon species and some populations will be able to withstand the changing circumstances and survive. Salmon biodiversity can also buffer their overall abundances from periods of environmental change, where some salmon traits are better adapted to different conditions than others. This has been referred to as the portfolio effect, where greater diversity is able to maintain more stable abundances to support harvest and ecosystem function (Schindler et al. 2010).

The WSP identifies implementation strategies that are designed to maintain diversity and abundances of Pacific Salmon to the fullest extent possible. But the WSP also acknowledges that there likely will be circumstances when losses of wild salmon populations are unavoidable. Some catastrophic events are beyond human control and DFO may not be able to restore habitat or spawning populations damaged by such events.

Pacific salmon are now facing significant challenges attributed to human activities, including accelerating global climate change, and others such as land-use, invasive species, disease and pollution, which are embedded within the overarching context of climate change (Grant et al. 2019). In response, Canadian Pacific salmon populations are generally exhibiting considerable declines in abundances and productivity, and consequently harvest has been significantly reduced (Grant et al. 2019; Grant et al. 2021;NPAFC 2019).

The rate of climate change in an area may exceed the ability of some salmon populations to adjust. Recent studies suggests that northward range contractions are likely for Pacific Salmon, and that some populations will be more vulnerable than others to local extinctions (Crozier et al. 2019, 2021; Cheung and Frölicher 2020). Although the intent of the WSP is to minimize biodiversity loss through fisheries or habitat degradation, where possible, it is unrealistic to expect that all losses can be avoided in natural environments.

This is consistent with expectations for Atlantic diadromous species in eastern North America. Diadromous species like Pacific and Atlantic salmon are more at risk to climate change given their complex life-cycles that rely on freshwater and marine habitats where impacts are cumulative. Atlantic diadromous species, which include Atlantic salmon, had the highest proportion vulnerable to climate change across 82 fish and benthic invertebrate species assessed in a recent climate change vulnerability assessment (Hayes and Kocik 2014; Hare et al. 2016).

Tracking Pacific salmon biodiversity is important for evaluating salmon responses to changing environmental conditions, to help identify and prioritize management actions when combined with other biological and socio-economic information, and also to evaluate the success of actions taken to reduce biodiversity and abundance losses. DFO's WSP Strategy 1 outlines a broad approach to tracking salmon biodiversity through standardized monitoring of wild salmon status. Within Strategy 1, the WSP prescribes three actions steps:

- Action Step 1.1.: Identify Conservation Units (CUs)
- Action Step 1.2: Develop Criteria to Assess CUs and Identify Benchmarks to Represent Biological Status
- Action Step 1.3: Monitor and Assess CU status

A.2 WSP ACTION STEP 1.1: IDENTIFY CONSERVATION UNITS

The Conservation Unit (CU) has been identified as the fundamental unit of biodiversity in the WSP (Strategy 1, Action Step 1.1). A CU is defined as a group of wild salmon sufficiently isolated from other groups that, if lost, are very unlikely to recolonize naturally within an acceptable timeframe (e.g. a human lifetime or a specified number of salmon generations) (DFO 2005).

Conservation Units are delineated by their salmon ecology, life-history, and genetics:

- Ecologies, specifically, ecological adaptive zones, are contiguous habitats that share common environmental and biological traits that shape the salmon adaptive environment. A Freshwater Adaptive Zone (FAZ) groups similar and contiguous freshwater habitats together based on their connectivity, climate, gradient, and hydrological characteristics. Marine Adaptive Zones (MAZ) groups similar and contiguous areas in the ocean that include coastal discontinuities within fjords, straits, and areas with distinct production processes, such as areas of upwelling and downwelling. The intersection of FAZ and MAZ are Joint Adaptive Zones (JAZ), which capture the full adaptive environment of Pacific salmon across their entire life-history. All salmon populations that fall within a JAZ are considered a potential CU, with the exception of sockeye that have a juvenile lake-rearing stage. This means that (in general) each JAZ contains at least one CU. These FAZ, MAZ and JAZ are all mapped geographically in BC/Yukon and Northeast Pacific Ocean.
- The next step in the process of delineating CUs is the refinement of geographic boundaries on a species-by-species basis based on the life history variation and genetics of each population.
- Life histories are phenotypic characteristics of salmon. Specific characteristics that further delineate CUs include adult-spawning run timing; coast versus interior spawners; and age of maturity. Sockeye are grouped into lake- versus river-type life histories. Lake-type are populations that rear in lakes for one to two years following their emergence from their spawning gravel; river-type are those that do not rear in freshwater, but shortly after they emerge from their spawning gravel they migrate downstream to the ocean, respectively. For lake-type sockeye, each lake or connected lake group can comprise a single CU. Therefore, sockeye have the largest number of CUs in BC and the Yukon, relative to Chinook, coho, chum and coho.
- CUs identified through their JAZ are further refined through the genetic structure of salmon species. In situations where the genetic analyses suggested significant population structure within a JAZ, life-history traits were examined to determine if the genetic structure corresponded to ecological patterns to further refine CUs.

Based on these considerations, there are 377 current CUs defined across the five species of salmon managed by DFO in the Pacific Region (Wade et al. 2019). Current is a specific CU designation. Current CUs are extant (i.e. not extirpated), and are either part of the original CU list developed by Holtby and Ciruna (2007), or have been modified and approved by CU experts and methodological processes that align with Holtby and Ciruna (2007). Revisions to the original CU list have been completed for Fraser sockeye and Southern BC Chinook (Grant et al. 2011; Brown et al. 2019). To ensure regional consistency in CU delineations and CU data sets, revisions to the CU list now must now be submitted to DFO Pacific Science's Data Management Unit for approval (Wade et al. 2019; Accessed June 1 2022: https://open.canada.ca/data/en/dataset/c48669a3-045b-400d-b730-48aafe8c5ee6/resource/bb3f949c-7f6d-4992-baf8-ec1c4d8c33ba).

A.3 WSP ACTION STEP 1.2: DEVELOP CRITERIA TO ASSESS CUS AND IDENTIFY BENCHMARKS TO REPRESENT BIOLOGICAL STATUS

DFO's WSP provides a framework for status assessments in its *Strategy 1, Action Step 1.2*. This includes a description of three status zones ranging from *Red* (poor status), to *Amber* (intermediate status), and *Green* (good status) (Figure 1; Table 1), and provides general methods for assessing a CU's status.

More detailed work was subsequently conducted to provide a CU status assessment toolkit (Holt 2009, 2010; Holt et al. 2009). To assess CU statuses, four classes of indicators were presented including abundance, trends in abundance, fishing mortality, and distribution (Figure 2).

To assess a CU's status, metrics are selected and applied depending on the availability of data and other information specific to each CU. Metric values are calculated and compared to the specific metric benchmarks to determine metric statuses (*Red*, *Amber*, or *Green*). For each metric, lower and upper benchmark values delineate, respectively, the *Red* to *Amber* and the *Amber* to *Green* WSP biological status zones (Figure 2).

Details on the metrics applied in the four completed WSP status assessments, and used in this analysis (Section 2.2.2).

A.4 WSP ACTION STEP 1.3: MONITOR AND ASSESS CU STATUS

Strategy 1 Action Step 1.3 is to monitor and assess statuses of CUs. The status assessment toolkit (Figure 2; Holt et al. 2009; Holt 2009; Holt 2010) was first applied to assess the statuses of Fraser sockeye CUs (Grant et al. 2011). This included identifying the appropriate metrics to use for these CUs, and also required considerable work to select and gap fill data for analyses. However, application of the metric toolkit did not result in completed status assessments for Fraser sockeye CUs, but instead flagged the need for subsequent work. Specifically, there were many Fraser sockeye CUs where statuses differed across individual metrics. For example, a CU's *percent change* (recent three generation) trend metric might indicate a *Red* status, while the *long-term trend* metric (comparing recent generation escapement average to long term average) might indicate a *Green* status. Given the different statuses that can be present across individual metrics for a single CU, simply applying the metric toolkit does not always produce sufficient advice on CU status. Therefore, an integration process that combines statuses across metrics into a single status was identified as a necessary next step, and was developed and implemented for Fraser sockeye CUs (Figure 2; Grant & Pestal, 2012).

WSP integrated statuses for the 24 Fraser sockeye CUs were produced through a Delphi-approach. This process used expert judgement from a group of salmon population specialists to combine status information across CU metrics, supplemented with additional information on CU productivity, abundance, and exploitation over time (Figure 3).

A key step for this status integration process is producing standardized data summaries for each CU, with status and other information to be used by the salmon experts to develop integrated statuses (Grant & Pestal 2012). These data summaries were essential for consistency, because they allowed all participants to work through the same quality-controlled information, rather than working off memory or different versions of data sets held by various individuals (reprinted in Appendix B).

The expert-driven process took three days and ~40 participants to produce statuses for 24 CUs, accompanied by narratives describing which information drove each status determination. Through this process, two additional status zones that fall between the original *Red*, *Amber*, and *Amber* zones were identified, these being *Red*/*Amber*, and *Amber*/*Green* statuses (Figure 3).

The WSP integrated status assessment approach, which includes applying the toolkit and integrating information to produce integrated statuses, was subsequently adapted and applied to Interior Fraser coho (DFO, 2015) and southern BC Chinook (DFO, 2016). This required considerable additional work in the data preparation stages for both of these groups of CUs to separate hatchery from wild populations prior to assessment. The status integration process for Southern BC Chinook was similarly laborious and time-intensive (e.g. Figure 4) as the first Fraser sockeye assessment, given the large number of CUs that were assessed. Interior Fraser coho was a shorter process (~25 people and 1 day), as there are only five CUs in this group.

Teams comprised of 4-6 individuals developed integrated status assessments for each CU, and then in plenary sessions, consensus on the final statuses were determined as an entire group. It is important to emphasize that there were different designations of status that varied among individuals in groups, and also between groups (Figure 3, Figure 4). Generally data that were absolute abundance and could be compared to existing relative-abundance benchmarks, provided more consistent assessments among participants. This was particularly the case when statuses were consistent across the probability levels of the estimated benchmarks; for example, *Red* statuses from the 10% to 90% probability level of the benchmarks, were likely to result in more consistent assignment of *Red* status among individuals and groups (Grant & Pestal 2012). Data that were indices of abundances only, where only trend metrics could be evaluated, led to more inconsistencies in status assignments among individuals and groups. This is also particularly the case when statuses differed between metrics.

The time investment to produce detailed WSP status assessments is significant, requiring the participation of 10-40 Pacific salmon experts for up to three days in a workshop setting. Further, after the workshops have been completed, the final status results proceed through CSAS peer-review processes, which require another set of formal meetings covering one or more days and 20-40 participants. In addition, effort is required to prepare data, calculate metric statuses, produce data summaries, organize workshops, and write and edit CSAS research documents. The resulting processes take years to complete, and for Interior Fraser coho, and Southern BC Chinook, CSAS research documents have yet to be published, to supplement the short CSAS Science Advisory Reports (DFO 2015; DFO 2016).

To re-assess the status of previously assessed CUs, a stream-lined status assessment process was developed and implemented for Fraser sockeye (Grant et al. 2020). The data summary packages were updated with more recent data to capture changes since the previous assessment. Data summaries were then run through a status integration process that involved only 10 Fraser sockeye experts over one day. However, even the stream-lined process took considerable work to prepare for and implement, and as a result, no re-assessments have occurred for Interior Fraser coho or Southern BC Chinook, and no further Fraser sockeye status re-assessments have occurred. Status assessments have not been completed, or even attempted for other species or areas.

All previous detailed WSP status assessment publications flagged the need for rapid annual status assessments, and guidelines on how often detailed reassessments should be required (DFO 2012; Grant & Pestal 2012; DFO 2015; DFO 2016). In the first Fraser sockeye WSP status process, a number of CUs were assigned provisional statuses, where it was recommended they be assessed regularly given concerns about their declining trends and anticipated rapid deterioration in status. The Interior coho process also recommended annual monitoring and status reassessment of CUs where there are signs of productivity and spawner abundance changes (DFO 2015). It was noted for this CU group that status metric values varied considerably over short time periods. For similar reasons, the Southern BC Chinook process also recommended annual assessments of status across individual metrics, the development of guidelines to integrate statuses across metrics, and more streamlined rapid assessment approaches going forward (DFO 2016).

The need for streamlined, rapid status assessments is particularly strong now, given the rapid declines being observed in salmon abundances and productivity, and the accelerating global climate change that is occurring (Bush and Lemmen 2019; Grant et al. 2019; Irvine et al. 2019; Crozier et al. 2019, 2021). Pacific salmon status assessments are likely to become rapidly out of date if not assessed frequently with these shifting conditions. The current inability to provide up-to-date CU statuses, and monitor statuses over time remains a key gap in implementing the WSP, specifically Strategy 1 Action Step 1.3: *Monitor and Assess CU status.*

Detailed WSP status processes continue to have merit in terms of fine-tuning data treatment, selecting metrics and estimating up-to-date benchmarks that may include time-varying productivity and/or changes to carrying capacity. Rapid status results can be used to flag and prioritize where more detailed integrated status assessments might be required to improve the rapid status algorithm going forward.

A.5 WSP VERSUS COSEWIC STATUS ASSESSMENTS

There are substantial similarities and some key differences between status assessments generated through DFO's WSP versus the COSEWIC. Similarities between WSP status assessments and COSEWIC include:

- WSP integrated status assessment methods were developed using COSEWIC criteria, which were based on IUCN status criteria, as a starting point;
- Both approaches rely on *percent change* and *absolute abundance* metrics, which are COSEWIC criteria that are also informally used for WSP status metrics;

- Recent COSEWIC assessments of Pacific salmon relied on the detailed WSP status materials (assembled data, WSP statuses, CU narratives) as the starting point for their assessments.

Specific differences between WSP and COSEWIC metrics used to assess stat (us include:

- The WSP *percent change* metric has a slightly more conservative lower benchmark of a 25% decline in abundance, compared to the COSEWIC benchmark at 30% decline (Holt et al. 2000; Grant et al. 2011).
- COSEWIC does not use a *relative-abundance* metric, despite its importance to the WSP status process. This is largely attributed to the fact that COSEWIC criteria are generic, designed to be applicable across a broad range of Canadian plant and animal species, while DFO's WSP status assessments are specifically designed for the nuances of Pacific Salmon, adapting existing COSEWIC and IUCN methods for this purpose. However, when relative-abundance benchmarks based on stock-recruitment data are not available for CUs, then the DFO and COSEWIC status assessment processes rely on similar metrics: abundance trends and absolute abundances if available.

Although COSEWIC includes a broad range of plant and animal species experts to assess status, they rely on core information provided by species experts. This contributed to results that are comparable between COSEWIC and DFO WSP status assessments, despite differences in their status assessment approaches.

The status assessment reports written by COSEWIC relied on data and information managed by DFO, collated from stock assessment projects by DFO, Indigenous groups and others. They also relied on DFO WSP CU designations, and status assessment results produced prior to the COSEWIC status process using identical data sets (Grant et al. 2011, 2020; Grant and Pestal 2012; DFO 2015, 2016).

There are slight differences in COSEWIC's definition of DUs, compared to DFO's WSP CUs, though COSEWIC largely relied on DFO's detailed CU methods (Holtby & Ciruna 2007) and results to identify DUs (Brown et al. 2019; Grant et al., 2011). Therefore the DU versus CU lists are identical for Fraser sockeye, and Interior Fraser coho, and have slight differences for a few Southern BC Chinook CUs.

There is also much overlap between COSEWIC's status assessments compared to DFO's WSP status assessments:

- Approximately 90% of CUs placed in the WSP *Red* or *Red/Amber* status zones aligned with COSEWIC's poorest status zone of Endangered (DU is facing an imminent threat of extinction).
- At the other end of the status spectrum, ~80% of CUs assessed in the WSP's healthiest status zone (*Amber/Green* and *Green*) aligned with COSEWIC's healthiest status zone of Not at Risk.
- In 80% of cases where DFO statuses were *Amber* (WSP intermediate status zones), the COSEWIC statuses were Threatened or Special Concern (WSP intermediate status zones between Endangered and Not at Risk).
- Notable differences occurred where a WSP status was more conservative, indicating a poorer status than COSEWIC (2 *Red*=2 Threatened; 1 *Amber*=1 Not at Risk). However, there was equal number of cases where WSP status was less conservative than COSEWIC (1 *Green*=1 Threatened; 2 *Amber/Green*=2 Special Concern).

APPENDIX B: STATUS NARRATIVE FROM INTEGRATED WSP STATUS ASSESSMENT WORKSHOPS

B.1 FRASER SOCKEYE

These status narratives were extracted from two sources:

- For the initial status assessments, with data up to 2010, we briefly summarized the detailed status narratives in Appendix 2 of Grant and Pestal (2012).
- For the 2017 reassessments, with data up to 2015, we extracted the very brief wording in Table 3 of the CSAS Science Advisory Report generated from the WSP status assessment (DFO 2018).

B.1.1 Early Stuart (SEL-06-14, *Red* in 2010, *Red* in 2015) (CYCLIC)

Stock Match: Early Stuart; this is considered a cyclic CU

The main considerations in the initial 2010 assessment of *Red* were: (1) set aside the *relative abundance* metrics due to concerns regarding S_{gen} and S_{msy} estimates for this highly cyclic CU, (2) steep decrease identified by *percent change* metric (*Red*), (3) *long-term trend* was *Amber*, but it fell close to the lower benchmark so did not alter the *Red* status designation.

The main considerations in the 2015 reassessment of *Red* were: (1) All metrics were *Red*, and in addition productivity was declining.

B.1.2 Chilliwack-ES (SEL-03-01, *Red/Amber* in 2010, *Amber/Green* in 2015)(CYCLIC)

Stock Match: Miscellaneous Early Summers; this is considered a cyclic CU

The main considerations in the initial 2010 assessment of *Red/Amber* were: (1) differences in integrated status determination between expert groups, and therefore, to the final mixed *Red/Amber* designation, was due to different interpretations of the same limited information for this CU particularly related to whether or not this CU was cyclic; (2) factors pointing to *Amber* designation were the *Amber* on the *relative abundance* metric if the arithmetic generational average is used and the absolute number of spawners well above the COSEWIC criterion D1 (assuming cyclic population dynamics). (3) The main factor pointing to *Red* designation was the *Red* status for the *relative abundance* metric using the geometric average of recent abundances; in addition, there are some recent years where abundances fall close to the COSEWIC criterion D1, when comparing all recent escapement data (assuming non-cyclic population dynamics).

The main considerations in the 2015 reassessment of *Amber/Green* were: (1) Rel Abd is *Amber*, (2) *percent change* and *long-term trends* are *Amber*, and (3) no years in the time series fell below the COSEWIC Criterion for small populations (1,000).

B.1.3 Pitt-ES (SEL-03-05, *Amber/Green* in 2010, Green in 2015)

Stock Match: Pitt

The main considerations in the initial 2010 assessment of *Amber/Green* were: (1) the mixed signals amongst metrics and status information presented in the data summaries, and the different interpretations of these mixed status signals between expert groups; (2) *relative abundance* metric statuses was *Green* at the 50% probability (median) benchmark for most models, but *Amber* at the adjacent higher probability level (75%); (3) systematic decreases in productivity with some recent years of productivity falling below replacement; (3) hatchery influence could be confounding the productivity time series, making productivity appear better than it actually is.

The main considerations in the 2015 reassessment of *Green* were: (1) all variations of the *relative abundance* metric were *Green*, (2) *long-term trend* was *Green* (3) *percent change* was *Red*, but has switched between *Red*, *Amber*, and *Green* several times over the time series, and 3 generation window included the largest abundances in the time series from the early 2000s.

B.1.4 Nahatlatch-ES (SEL-05-02, *Red* in 2010, Amber in 2015)

Stock Match: Miscellaneous Early Summers

The main considerations in the initial 2010 assessment of *Red* were: (1) *Red* for the *percent change* metric, (2) some recent low years of abundance that fall below 1,000, (3) *long-term trend* was *Amber*, but very close to the lower benchmark.

The main considerations in the 2015 reassessment of *Amber* were: (1) low absolute abundance (median 2000, 1 of last 4 years < 1,000), (2) *long-term trend* and *percent change* were *Amber*.

B.1.5 Anderson-Seton-ES (SEL-06-01, Amber in 2010, *Amber/Green* in 2015)

Stock Match: Gates

The main considerations in the initial 2010 assessment of *Amber* were: (1) This is a cyclic CUs, so the *relative abundance* metric was not considered at the time, (2) overall population increase since the 1960s and 1970s, resulting in *Amber* for the *long-term trend* metric, (3) stable abundance in recent years, (4) recent declining trend, resulting in *Red* for the *percent change* metric, (5) relatively low abundance on weak cycle years; however, no recent years fall below 1,000.

The main considerations in the 2015 reassessment of *Amber/Green* were: (1) *percent change* and *long-term trend* were *Green*, (2) no years out of the last four fell below the COSEWIC Criterion for small populations (1,000), (3) the general declining productivity pattern for this CU contributed to lowering the status to *Amber/Green*, although there have been improvements in productivity since the 2012 status assessment, (4) mixed *Amber*/*Green* on the relative-abundance metric (i.e. at probability levels below 75% this metric was *Amber* and at probability levels above the 50% probability level this metric was *Amber*).

B.1.6 Taseko-ES (SEL-06-16, *Red** in 2010, *Red* in 2015)

Stock Match: Miscellaneous Early Summers

The main considerations in the initial 2010 assessment of *Red* were: (1) consistently *Red* status for trend metric (percent change and long-term), (2) this CU does not have recruitment data, therefore, *relative abundance* metric statuses could not be estimated, (3) since abundance data for this CU are an index only, recent absolute abundances could not be compared to COSEWIC Criteria D1.

Overall, the integrated status in 2010 for this CU was flagged as provisional, because data quality was rated fair; Workshop participants highlighted that escapement data, which are an index of escapement only, would require further evaluation.

The main considerations in the 2015 reassessment of *Red* were: (1) *long-term trend* and *percent change Red*, (2) no recruitment estimates available, so no benchmarks for the *relative abundance* metric.

B.1.7 Nadina-Francois-ES (SEL-06-20, *Red* in 2010, *Amber/Green* in 2015)

Stock Match: Nadina

The main considerations in the initial 2010 assessment of *Red* were: (1) *Red* on the *relative abundance* metric status across 23 of 30 paired upper and lower benchmark combinations (probability levels and model forms); (2) systematic decreases in productivity, (3) *Red* on the *percent change metric* was discounted because the CU was returning from a periods of large abundance, particularly the year 2000. (4) *long-term trend* was *Green*, (5) a few expert groups assigned this CU a provisional *Red* status in the initial assessment due to concerns regarding SR model fit and resulting benchmark estimates.

The main considerations in the 2015 reassessment of *Red* were: (1) *Relative abundance* metrics was *Amber* at the median benchmark estimate, and *Red* above, (2) long-term and *percent change* trends were both *Amber*.

B.1.8 Bowron-ES (SEL-07-01, *Red* in 2010, *Red* in 2015)

Stock Match: Bowron

The main considerations in the initial 2010 assessment of *Red* were: (1) all metrics were *Red*.

The main considerations in the 2015 reassessment of *Red* were: (1) all metrics were *Red*.

B.1.9 Shuswap ES (SEL-09-02, *Amber/Green* in 2010, Amber in 2015)(CYCLIC)

Stock Match: Scotch, Seymour, Miscellaneous Early Summer. This is considered a cyclic CU.

The main considerations in the initial 2010 assessment of *Amber/Green* were: (1) *relative abundance* metric not considered due to concerns regarding the appropriate for cyclic CUs, (2) large and building abundances on the dominant cycle (*Green long-term trend* metric) pointed to *Green* status, (3) increasing productivity in recent years pointed to *Green* status, (4) concerns over *Red* on *percent change* metric (5) one very recent observation of low abundance on a weak cycle that falls below the COSEWIC Criteria D1 of 1,000, and recent decreases in abundance in the off-cycle years pointed to *Amber* or *Red status*, however, most of this decreasing trend was attributed to a single weak cycle year in 2009, therefore, this decreasing trend alone was not sufficient to place this CU in a *Red* status zone.

The main considerations in the 2015 reassessment of *Amber* were: (1) *Amber* on the *relative abundance* metric on the dominant cycle line if cycle-specific benchmarks from the Larkin model are considered, (2) although the three other cycles are *Red*, they did not drive the integrated status of this CU, (3) consistently *Green* statuses across the *long-term trend* and *percent change* metrics.

B.1.10 Kamloops-ES (SEL-10-01, Amber in 2010, Amber in 2015)

Stock Match: Raft, Miscellaneous Early Summer

The main considerations in the initial 2010 assessment of *Amber* were: (1) *Amber* on relative-abundance metric paired upper and lower benchmark combinations at the median probability levels (50%) for all models but the recursive Bayesian Ricker model; however, since this CU does not exhibit any systematic productivity trends, models that consider recent productivity (such as the recursive-Bayesian Ricker model) were not given high weight in relative-abundance metric status evaluations; (2) *Green* on *long-term trend* metric, which provides extra weight to the *relative-abundance* metric status, which were mostly *Amber,* with some *Reds* at higher probability levels, (3) *percent change* metric was *Red*, but down-weighted given this CU was returning from a period of high abundance and was not exhibiting any systematic trends in productivity.

The main considerations in the 2015 reassessment of *Amber* were: (1) *Amber* on *relative abundance* metric, but with high uncertainty in the benchmark estimates, (2) *Green* on *long-term trend*, (3) *Red* on *percent change*, but after coming off a period of high production in the mid-1990’s.

B.1.11 North-Barriere-ES (SEL-10-03, Amber in 2010, Amber in 2015)

Stock Match: Fennel, Miscellaneous Early Summer

The main considerations in the initial 2010 assessment of *Amber* were: (1) *Amber* statuses across 29 of 30 paired upper and lower benchmark combinations (probability levels and model forms); however, the lower benchmarks were flagged as being low, ranging from 300 to 3,000, depending on model form and probability level) relative to the COSEWIC Criteria D1 values of 1,000. (2) very recent productivity appeared to be stable or increasing, (3) although *percent change* was *Red*, this metric was down-weighted given this CU was coming off a period of higher abundances, (4) *Green long-term trend* status, although some groups felt *long-term trends* should be given lower weight due to the higher exploitation rates in earlier years.

The main considerations in the 2015 reassessment of *Amber* were: (1) *Amber* on the *relative abundance* metrics, but noting that the lower benchmark (S_{gen}) at 1,000 was the same as the *Endangered* threshold used by COSEWIC, (3) *Green* on *long-term trend* metric, (3) *Red* on *percent change* metric.

B.1.12 Takla-Trem-S-S (SEL-06-13, *Red/Amber* in 2010, *Red/Amber* in 2015)(CYCLIC)

Stock Match: Late Stuart. This is considered a cyclic CU.

The main considerations in the initial 2010 assessment of *Red/Amber* were: (1) Highly cyclic, so *relative abundance* metric was set aside (concerns of S_{gen} and S_{msy} estimates for cyclic CUs), (2) factors pointing to *Amber* included large absolute abundance (no recent abundances near the COSEWIC Criteria D1 for small populations), *Green long-term trend* metric, and *Red percent change* metric (however, this CU is returning to average following a period of high abundance), (3) factors pointing to *Red* included decreasing productivity combined with (including a few recent years of less than 1 R/S) in combination with steep decline in *percent change* (although it was noted that this CU is coming off a period of high abundance, the steepness of the recent decrease in abundance was flagged as a concern).

The main considerations in the 2015 reassessment of *Red/Amber* were: (1) mixed *Red* and *Amber* on *relative abundance* metric across p-levels and cycle lines if the Larkin-based benchmarks are considered, (2) *Red* on *percent change* metric, (3) productivity has declined and remained low for the past 15 years (three of these years exhibited below replacement productivity), (4) continued to decline in abundance since the previous assessment (4) the dominant cycle line and one weak cycle line have decreased in abundance, and the subdominant and second weak cycle lines have respectively, remained stable and increased (5) *long-term trend* status is *Green* and has been *Green* for most of the retrospective metric evaluation, but this metric alone was not sufficient to bump up the integrated status from *Red/Amber*.

B.1.13 Quesnel_S (SEL-06-10, *Red/Amber* in 2010, *Red/Amber* in 2015)(CYCLIC)

Stock Match: Quesnel. This is considered a cyclic CU.

The main considerations in the initial 2010 assessment of *Red/Amber* were: (1) Highly cyclic, so *relative abundance* metric was set aside (concerns of S_{gen} and S_{msy} estimates for cyclic CUs), (2) decreasing productivity pointed to *Red*, with a several years below 1 R/S, (3) *Red* on *percent change* metric, with concerns due to a steep decline while noting that this CU was returning to average abundances after a period of high abundance, (4) large absolute abundance pointed to *Amber*, (5) concerns regarding the estimated productivity trends using Ricker model residuals, which may not be capturing the effects of cycle-line interactions on the productivity trends (Larkin model may be more appropriate), (6) *Green* on *long-term trend* metric but disregarded because this CU's early time series was low after a period of human activities that significantly reduced this population's size and the long-term time series does not provide appropriate comparison for the *long-term trend* metric.

The main considerations in the 2015 reassessment of *Red/Amber* were: (1) mixed statuses for the relative-abundance metric, showing *Red* on the dominant cycle at probability levels below 50%, and *Amber* above, *Amber* on the subdominant cycle, and *Red* on the two weak cycles; it appears that the dominant cycle might be shifting, with the new dominant cycle emerging on the previously subdominant cycle, (2) declines in productivity in the recent decade pointed to *Red*, (3) *Red* on *percent change*, (4) abundance was relatively low from 2006 to 2013 and one year in the last four falls below the COSEWIC Criterion for small populations (1,000), (5) positive signals in the slight increase in the R/ETS time series in recent years, though the Larkin model residuals do not indicate a similar increase (5) *Green* status of the *long-term trend* metric (which has been *Green* for the entire time series).

B.1.14 Chilko-S-ES (SEL-06-02, Green* in 2010, Green in 2015)

Stock Match: Chilko

The main considerations in the initial 2010 assessment of *Green* were: (1) relative-abundance metric *Green* across all benchmark probability levels and model forms, with high data quality, (2) *Red* on the *percent change* metric down-weighted, because relative and absolute abundance were large and CU returning to average abundance following a previous period of high abundance, (3) in very recent years both abundance and productivity have increased (4) Integrated *Green* status was flagged as provisional, given the potential status deterioration in the *percent change* if recent productivity (recruits/spawner) and abundance trends persist, (5) a few recent years of less than 1 R/S, but this could be linked to high spawner abundance (density dependence).

The main considerations in the 2015 reassessment of *Green* were: (1) relative-abundance metric *Green* across all benchmark probability levels and model forms, with high data quality, (2) *Green* on *long-term trend*, (3) productivity and *percent change* (*Green*) have improved; in the previous assessment both these trends were declining and were flagged as a risk to deteriorating status had they continued.

Note that Chilko-ES is distinct from the Chilko-S CU (different run timing and spawning locations in the Chilko watershed), but the data for Chilko-ES currently has not been disaggregated from the larger Chilko-S CU; the Chilko-ES abundance comprises less than 10% of the combined Chilko-S & Chilko-ES aggregate. Integrated status could not be evaluated for Chilko-ES for either 2010 or 2015 given there are no independent data available. In the 2012 workshop looking at data up to 2010, participants recommended that an escapement index and proxy exploitation rate for the Chilko-ES CU be developed to provide information for subsequent status evaluations.

B.1.15 Fran-Fras-S (SEL-06-07, *Red/Amber* in 2010, *Amber/Green* in 2015)

Stock Match: Stellako

The main considerations in the initial 2010 assessment of *Red/Amber* were: (1) most participants agreed on a provisional *Amber* integrated status designation for this CU, but due to differences both within and amongst expert groups, this CU was designated a blended *Red/Amber* status, (2) factors pointing to *Amber* included high recent abundance and *Green long-term trend* while down-weighting the *Red percent change*, because abundances for this CU are returning to average following a previous period of above-average abundance, (3) factors pointing to *Red* included recent declines in CU productivity, with some years falling below replacement, and *Red* status for the relative-abundance metric when looking at the time-varying probability model fit, and a decreasing *percent change*.

The main considerations in the 2015 reassessment of *Amber/Green* were: (1) mixed *Amber*/*Green* status for the *relative-abundance* metric, depending on probability level, (2) *Green* on *long-term trend*, and *Amber* on *percent change*.

B.1.16 Cultus-L (SEL-03-02, *Red* in 2010, *Red* in 2015)

Stock Match: Cultus

The main considerations in the initial 2010 assessment of *Red* were: (1) *Red* on relative-abundance metric across probability levels and model forms (2) *Red* on *long-term* and *percent change* metric, (3) recent abundance below 1,000, the COSEWIC Criterion D1 for small populations, (4) productivity trends have also decreased in recent years.

The main considerations in the 2015 reassessment of *Red* were: (1) *Red* on relative-abundance metric across probability levels and model forms, (2) *Red long-term trend*, (3) for three of the last four years and nine of the last 12 years effective female spawners fell below the COSEWIC Criterion for small populations (1,000); (4) productivity decreased in years preceding the 2000 brood year, before hatchery intervention; productivity data after 2000 could not be compared since these values are confounded by hatchery intervention; (5) *percent change* metric is *Amber*, the slightly increasing abundance is being supported by hatchery intervention (second generation hatchery enhanced fish that are unmarked); therefore, this metric is not given much weight; it is unclear whether this CU would currently be sustainable in the absence of hatchery intervention.

B.1.17 Harrison-DS-L (SEL-03-03, Green in 2010, *Amber/Green* in 2015)

Stock Match: Miscellaneous Late

The main considerations in the initial 2010 assessment of *Green* were: (1) *Green* on *percent change* and *long-term trends*, (2) this CU does not have recruitment data, therefore, *relative abundance* metric statuses could not be estimated, (3) absolute abundance cannot be directly compared to COSEWIC Criteria for this CU since only one out of a number of creeks is being used as an indicator of this CU's status; however, for this single creek alone (Big Silver), it does not trigger COSEWIC's Criteria D1 in the last four years, (4) although the *percent change* in abundance metric was *Green* in status, in very recent years there has been a decrease in abundance and it was recommended that this trend be monitored, since if it persists the status of this CU could change in the near future (to *Amber* or *Red*).

The main considerations in the 2015 reassessment of *Amber/Green* were: (1) *Green long-term trend*, (2) *Red percent change*, but coming off a peak abundance, (2) *relative abundance* metric not available.

B.1.18 Harrison-US-L (SEL-03-04, Amber in 2010, *Red* in 2015)

Stock Match: Weaver

The main considerations in the initial 2010 assessment of *Amber* were: (1) *relative abundance* metric mostly *Amber* across benchmark probability levels and model forms, (2) *long-term trend* was also *Amber*, (3) *percent change* was not weighted high given absolute abundance was high, (4) recommended frequent monitoring of the *percent change* (which was *Red*, because it could produce changes in other metric statuses, and therefore, integrated status, if this trend persists.

The main considerations in the 2015 reassessment of *Red* were: (1) All metrics were *Red*, (2) Two of last four 4 years has less than 1,000 spawners

B.1.19 Lillooet-Harrison-Late (SEL-04-01, Green* in 2010, Amber in 2015)

Stock Match: Birkenhead

The main considerations in the initial 2010 assessment of provisional *Green* were: (1) absolute abundance well above COSEWIC Criteria D1 on small populations for the entire time series, even though the time series only covers Birkenhead River, so absolute abundance for the CU is, in fact, higher than indicated by the data, (2) *percent change* and *long-term trend* are *Green*, (3) relative-abundance metric changed to *Amber* status only in recent years (4) designated a provisional *Green* integrated status, given the declining productivity trends

The main considerations in the 2015 reassessment of *Amber* were: (1) Rel Abd metric is *Amber*, (2) *long-term trend* metric is *Green,* (3) *percent change* is *Red*, (4) low productivity trend, but combined with high absolute abundance.

B.1.20 Seton-L (SEL-06-11, Undetermined in 2010, *Red* in 2015)

Stock Match: Portage

The main considerations in the initial 2010 assessment of Undetermined were: (1) *Relative abundance* metrics set aside, due to concerns regarding benchmark estimates for cyclic CUs, (2) no integrated status designation could be agreed upon by workshop participants; the integrated status designated by groups included all three WSP status zones (*Red*, *Amber*, and *Amber*); even within groups there was inconsistency in status determinations amongst individuals:

- two groups designated this CU *Red* based on the steep decline in abundance and (*percent change*) and the decreasing productivity,
- two groups designated this CU *Green*, emphasizing that the dominant cycle did not exhibit any decreasing trend in abundance and has been quite stable since the 1980's (after a period of rebuilding in the previous decade after the original CU was extirpated); these groups discounted the *percent change* and *long-term trend* metric in their status evaluations since they felt these metrics were strongly influenced by a single low observation on a single subdominant cycle year;
- one group designated this CU *Amber*, as a middle ground to balance all the considerations presented by the *Red* and *Amber* designations described in previous bullets; although the group agreed to an *Amber* integrated status, interpretations varied amongst individuals in this group;

The main considerations in the 2015 reassessment of *Red* were: (1) *relative abundance* is *Red* across variations of model form and probability level, (2) Two of last four years had less than 1,000 spawners, (3) *long-term trend* and *percent change* have been *Red* for several years.

B.1.21 Shuswap-L (SEL-09-03, Green in 2010, *Amber/Green* in 2015) (CYCLIC)

Stock Match: Late Shuswap. This is considered a cyclic CU.

The main considerations in the initial 2010 assessment of *Green* were: (1) *relative abundance* metrics set aside, due to concerns regarding benchmark estimates for cyclic CUs, (2) *long-term trend Green* and generational average abundance increasing, (3) stable productivity, (4) large number of spawners on the dominant cycle year for this CU (last dom year was 5.5 million).

The main considerations in the 2015 reassessment of *Amber*/*Green* were: (1) *Green* status of the dominant cycle relative-abundance, (2) large number of spawners on the dominant cycle year for this CU (last gen avg 2.1 million), (3) productivity stable since the beginning of the time series with increase in very recent years, (4) the low abundances of the other cycle-lines, and declining trends observed for the subdominant cycle, down-weighted this CU's status to *Amber*/*Green*, (5) though the *percent change* trend status is *Red*, this is driven by the subdominant (which are largely five year old fish from the dominant cycle) and the first (and smallest) weak cycle; the dominant cycle has not exhibited a declining trend, and in fact had exceptional returns in the last two cycle years (2010 and 2014). Since the weak cycle is not enumerated with high precision methods (visual methods applied), a sensitivity analysis of the potential error in the recent weak cycle estimate indicates that the trend metric status could range from *Red* to *Amber*, depending on the true value, (6) similarly, the *long-term trend* of 0.46 is *Red* status, but falls right on the edge of an *Amber* status (lower benchmark is 0.5); if the most recent weak cycle abundance were actually greater than 100 (which is within the range of error for this data point), this metric status would change to *Amber*;

B.1.22 Widgeon-RT (SER-02, *Red* in 2010, *Red* in 2015)

Stock Match: Miscellaneous Late

The main considerations in the initial 2010 assessment of *Red* were: (1) low absolute abundance, falling below COSEWIC D1 for a number of recent years, (2) *Red* on *long-term trend* metric, (3) the current generational average abundance (89) is extremely small, (4) this CU does not have recruitment data, so no relative-abundance benchmarks, (5) *Amber* on *percent change* metric does provide some encouraging indications of improving trends, however, these trends were not sufficient to change this CU's integrated status designation from *Red*.

The main considerations in the 2015 reassessment of *Red* were: (1) absolute abundance low, with 3 of last 4 spawner abundances less than 1,000, (2) this CU has a small spatial distribution, therefore, it CU will be consistently in the *Red* status zone.

B.1.23 Harrison R (SER-03, Green in 2010, Green in 2015)

Stock Match: Harrison

The main considerations in the initial 201 assessment of *Green* were: (1) *relative abundance Green* at median benchmark estimates, (2) *percent change* and *long-term trend Green*, (3) productivity has also increased over the past decade, (4) average absolute abundance in the last generation has been an order of magnitude higher than the time series average.

The main considerations in the 2015 reassessment of *Green* were: (1) *relative abundance Green* at median benchmark estimates, (2) *percent change* and *long-term trend Green*.

B.2 SOUTHERN BC CHINOOK

These status narratives were extracted from Tables 9 to 14 of the unpublished Working Paper generated after the expert workshop, which used data up to 2012. Note that the status assessments only apply to the CU (i.e. are based on data for wild sites, and exclude data for sites with moderate or high levels of hatchery supplementation).

These short narratives document the considerations identified in plenary discussion as determining the status designation. Status notes were developed during plenary discussions on Day 2 and 3, first with CU names hidden, and then revealed. The summaries reproduced here were shortened to focus on the key drivers in the deliberation, but all additional comments raised by participants were merged into the CU summary in Appendix B of the unpublished Working Paper. At almost 200 pages, the detailed notes are too long to reproduce here.

CUs with integrated status designations are listed first, in sequence of CU ID. CUs for which no integrated status was assigned are then grouped together, summarizing the rationale provided.

Data deficient CUs were also grouped into five types, with details included for each CU below:

- Type 1: Time series of good quality data available, but considered not representative of whole CU.
- Type 2: Good quality data available, but time series too short to make inferences about trends.
- Type 3: Data available, but none meet the quality criteria
- Type 4: Good quality data available, but none for sites classified as wild.
- Type 5: No recent data

B.2.1 Okanagan 1.x (CK-01, *Red*)

Based on metrics (all *Red*) and very low relative index of abundance (peaks at 30 fish), this CU was classified as *Red*, but it is very likely extirpated. (CU definition should be revisited given the presence of US hatchery strays.)

B.2.2 Lower Fraser River FA 0.3 (CK-03, Green-provisional)

WSP metrics for *relative abundance* and extent of decline are *Green*, and absolute abundance is substantial for Chinook. However, the short-term decline observed in recent years, despite decreasing exploitation rates, resulted in a provisional status designation to highlight the need for monitoring the trend.

B.2.3 Lower Fraser River-Upper Pitt SU 1.3 (CK-05, Data Deficient – Type 1)

Based on available data and the metrics presented, most groups assessed this CU as *Red* due to declining trends and low abundance. However, participants agreed to a DD assessment based on additional information provided by a participating expert (the single site with data is not representative, and surveys of additional sites within the CU are currently not feasible). Specifically, the rationale was "*Time series of good quality data available, but considered not representative of whole CU. Only 1 population surveyed but others may exist that are not yet known.*"

B.2.4 Lower Fraser River SU 1.3 (CK-06, Data Deficient – Type 1)

Time series of good quality data available but considered not representative of whole CU. Data available for only 1 site out of 7 (most abundant site cannot be assessed due to low visibility), and for the site with data, the time series is too short.

B.2.5 Maria Slough SU 0.3 (CK-07 – TBD)

The CU has received an enormous amount of stewardship and watershed restoration activity. Human land-use impacts have changed the hydrography of this geographically small CU. There is no data for wild sites in the CU.

B.2.6 Middle Fraser-Fraser Canyon SP 1.3 (CK-08, Data Deficient – Type 3)

Data available, but none meet the quality criteria. Only records are opportunistic observations during sockeye salmon surveys.

B.2.7 Middle Fraser River-Portage FA 1.3 (CK-09, *Red*)

Most groups assessed this CU as *Red* status based on *relative abundance* (*Red* even if doubled the index spawners) and the *percent change* / probability of decline combination. However, there is high uncertainty due to short time series of data with quality ranking and observed lack of response to decreasing ER.

B.2.8 Middle Fraser River SP 1.3 (CK-10, *Red*)

Strong and significant downward trend. Even if true abundance were double the estimate due to bias in relative index, it would still not exceed S_{gen}.

B.2.9 Middle Fraser River SU 1.3 (CK-11, Amber)

All groups assessed this CU as *Amber* due to mixed signals from the 3 metrics (1 *Red*, 1 *Amber*, 1 *Green*). Overall, the magnitude of decline is not large, not all sites are declining and the total of index spawners is well above COSEWIC Criterion D. In combination, this resulted in a down-weighting of the *percent change* metric (*Red*).

B.2.10 Upper Fraser River SP 1.3 (CK-12, *Red*)

Relative abundance, *percent change*, and probability of decline are all *Red*. Very small contribution of hatchery in recent years (few hundred hatchery fish among tens of thousands of spawners), moderate precision, but highly reliable aerial survey estimates. Overall, highly confident in assessment.

B.2.11 South Thompson SU 0.3 (CK-13, Green)

Percent change and extent of decline show pronounced increase, and *relative abundance* metric should be green as well with a relatively small adjustment for likely under-estimate in relative index.

B.2.12 South Thompson SU 1.3 (CK-14 *Red/Amber*)

Participants settled on *Red/Amber* based on a show of hands after much debate within and between groups. Key considerations were: *relative abundance* and *percent change* are *Red*, but visual estimates are imprecise, may be biased low and fall near the confidence range for the lower benchmark, so might actually be *Amber* on the *relative abundance* metric. Also, anecdotal information was presented that 2013 had a large return (not included in the data summary) which further moved the evaluation towards *Amber*.

B.2.13 South Thompson-Bessette Creek SU 1.2 (CK-16, *Red**)

Precipitous decline and extremely low numbers (but need to revisit CU definition). If this is accepted as a CU, then no question that the population has declined drastically.

B.2.14 Lower Thompson SP 1.2 (CK-17, *Red*)

Most groups designated this CU as *Red* status based on metrics for *relative abundance* (*Red*) and *percent change* (*Red*), but 1 group leaned to *Amber* designation based on extent of decline (*Amber*) and *relative abundance* after rough adjustment for sites not surveyed (i.e. index spawner abundance close to lower benchmark).

B.2.15 North Thompson SP 1.3 (CK-18, *Red*)

Very strong short-term decline and very low numbers of fish, combined with high uncertainty due to small number of data points.

B.2.16 North Thompson SU 1.3 (CK-19, *Red*)

Despite being a relative index of abundance, the WSP benchmarks metric (*Red*) carried significant weight in this case because 4 out of 5 available sites are included in the data stream. *Percent change* (*Red*) and probability of decline (*Red*) were also strong indicators.

B.2.17 Southern Mainland-Georgia Strait FA 0.x (CK-20, Data Deficient – Type 5)

No recent, high quality escapement records.

B.2.18 East Vancouver Island-Nanaimo SP 1.x (CK-23, Data Deficient – Type 5)

Very little recent data exists. A genetic baseline review is currently ongoing to determine whether this CU still exists.

B.2.19 Southern Mainland-Southern Fjords FA 0.x (CK-28, Data Deficient – Type 2)

No biological benchmarks presented and available spawner time series is short. However, workshop participants highlighted that data for this CU need to be further investigated, and categorized it as "*Good quality data available, but time series too short to make inferences about trends*."

B.2.20 East Vancouver Island-North FA 0.x (CK-29, *Red*)

All groups designated this CU as *Red* but this was the result of considerations other than the 3 WSP metrics. Rather, participants highlighted the following concerns: only a small portion of total abundance in wild sites, impacts of straying are likely and a very small index of abundance of wild sites.

B.2.21 West Vancouver Island-South FA 0.x (CK-31, *Red*)

Most groups designated this CU as *Red*, but due mostly to pressures (straying from large-scale hatchery releases, including sea pens, and high exploitation rates (roughly 60%) rather than to abundance or observed trends. Data from 2 small populations among 21 possible wild sites is not considered to be representative. Participants recommended completion of further work to determine whether these populations still exist as a CU under WSP definition.

B.2.22 West Vancouver Island-Nootka & Kyuquot FA 0.x (CK-32 , *Red*)

Most groups designated this CU as *Red*, but this was the result of considerations other than the 3 WSP metrics. Rather, participants highlighted the following concerns: only a small portion of total abundance in wild sites and impacts of straying are likely, very small index of abundance of wild sites.

B.2.23 Homathko SU x.x (CK-34, Data Deficient – Type 5)

Very little recent data exists. Regular visual surveys are not feasible on this large and turbid river.

B.2.24 Klinaklini SU 1.3 (CK-35, Data Deficient – Type 5)

There has been no data collected in recent years, and no supporting information exists to inform a status evaluation at this time. Past fishwheel surveys showed large number of Chinook (7k to 18k), but no data from recent years and not part of regular survey program.

B.2.25 Upper Adams River SU x.x (CK-82, Data Deficient – Type 3)

Initial *Red/Amber* status designation was based on data presented, but participants agreed to a DD status based on additional information provided by a participating expert (the site data quality was misclassified). The available spawner estimates are based on redd counts, which are difficult to assess consistently, and the CU is not routinely surveyed.

B.2.26 Type-4 Data Deficient (Good quality data, but none for wild sites) – 11 CUs

Workshop participants initially attempted to assess the status of the enhanced units, but eventually agreed that the WSP status metrics and benchmarks cannot be applied directly and recommended further work on methods specifically for these cases.

For 10 of these 11 cases, the status assessment was To Be Determined (TBD), with the rationale that the CU was “Not assessed due to unresolved technical and policy aspects of evaluation approach. These CUs were, in order of ID:

- CK-21 - East Vancouver Island-Goldstream_FA_0.x
- CK-33 - West Vancouver Island-North_FA_0.x
- CK-22- East Vancouver Island-Cowichan & Koksilah_FA_0.x
- CK-02 - Boundary Bay_FA_0.3
- CK-07 - Maria Slough_SU_0.3
- CK-25 - East Vancouver Island-Nanaimo & Chemainus_FA_0.x
- CK-15 - Shuswap River_SU_0.3
- CK-83 - East Vancouver Island-Georgia Strait_SU_0.3
- CK-27 - East Vancouver Island-Qualicum & Puntledge_FA_0.x
- CK-9008 - Fraser-Harrison fall transplant_FA_0.3

One of these CUs, Lower Fraser River_SP_1.3 (CK-04) was flagged for a review of enhancement classification. Specifically, the rationale was: “Not assessed, given unresolved technical and policy aspects of evaluation approach. However, the classification of enhancement level needs to be reviewed because enhancement stopped in 2002 brood year and the system now has natural spawners. There are also a number of locations within this TU that have no enhancement but are not surveyed.

B.3 INTERIOR FRASER COHO

These status narratives were extracted from Table 4 in the Science Advisory Report generated from the WSP status assessment (DFO 2015).

B.3.1 Middle Fraser Coho (Amber)

The main considerations in the integrated status determination were: (1) the patterns of productivity with frequent failures to achieve replacement over the most recent 13 years, (2) the low productivity and low smolt-adult survival over the last two decades, (3) the poorly described and imprecise stock-recruitment function, (4) moderate to high level of uncertainty and variability for the information presented, and (5) the current spawner abundance relative to benchmark estimates and COSEWIC reference points.

B.3.2 Fraser Canyon Coho (Amber)

The main considerations in the integrated status determination were (1) the *percent change* was *Red* for all but the most recent year and there is a moderate probability that the CU is currently in the *Red* zone, (2) the patterns of productivity with frequent failures to achieve replacement over the most recent 13 years, (3) the low productivity and low smolt-adult survival since 1998, (4) the short time series with no information prior to 1998, (5) the abundance exceeded the COSEWIC reference points, (6) the CU has a small capacity and has low-moderate intrinsic productivity, and (7) this is a single-site CU with spawners in a short section of one river which reduces resilience to perturbations and there is no likelihood of replacement from adjacent tributaries.

B.3.3 Lower Thompson Coho (Amber/Green)

The main considerations in the assignment of mixed status were (1) the *percent change* was increasing and there was virtually 0% probability for the *Red* zone, (2) the extent of decline metric showed the recent spawner abundance was above the long-term average and generally above the average level during the period of higher productivity (pre 1991), (3) the last four years exceeded the upper abundance-based benchmark, (4) the patterns of productivity with frequent failures to achieve replacement over the most recent 13 years—with 3 of the last 6 years very near replacement, (5) the low productivity and low smolt-adult survival since 1998, and (6) the steadily increasing trend in smolt production since 1995.

B.3.4 North Thompson Coho (Amber/Green)

The main considerations in the integrated status determination of mixed status were (1) the *percent change* was increasing, (2) the extent of decline metric showed the recent two years had increased but it was in the yellow or *Red* zone in the eight previous years, (3) productivity was often below replacement (6 of the last 13 brood years), (4) spawner abundance exceeded the upper confidence limit for the upper benchmark over the last three years, and (5) smolt-adult survival has been low and stable since brood year 2000.

B.3.5 South Thompson Coho (Amber)

The main considerations in the integrated status determination were (1) the patterns of productivity with frequent failures to achieve replacement over the most recent 13 years, (2) the low productivity and low smolt-adult survival over the last two decades, (3) the poorly described and imprecise stock-recruitment function, (4) moderate level of uncertainty and variability with the information presented, and (5) the spawner abundance relative to benchmark estimates.

APPENDIX C: DATA USABILITY, METRICS, AND INTEGRATED STATUS ASSESSMENTS

This Appendix lists the WSP metrics and integrated status assessment results from the WSP processes. It also includes a summary of which metrics are applicable for each CU. For Southern BC Chinook CUs, this *data usability* summary was taken from Brown et al. (2020). For Fraser sockeye and Interior Fraser coho, we applied the same approach as Brown et al. (2020), based on the data notes in the WSP status documentation. The following species and areas have been covered:

- *Fraser River Sockeye* (Appendix C.1): Fraser River sockeye were formally assessed under the WSP in 2011 (Grant and Pestal 2012) and re-assessed in 2016 (Grant et al. 2020). Table 13 lists data usability by CU. Table 14 and Table 15 list the status metric values used at the time, and the resulting expert assessments of integrated status.
- *Southern BC Chinook* (Appendix C.2): Southern BC Chinook were formally assessed under the WSP in 2012 (DFO 2016). Table 16 and Table 17 list data usability by CU. Table 18 lists the status metric values used at the time, and the resulting expert assessments of integrated status.
- *Interior Fraser Coho* (Appendix C.3) Interior Fraser Coho were formally assessed under the WSP in 2012 (DFO 2015). Table 19 lists data usability by CU. Table 20 lists the status metric values used at the time, and the resulting expert assessments of integrated status.

C.1 FRASER SOCKEYE DATA USABILITY, METRIC VALUE, AND INTEGRATED STATUS ASSESSMENT RESULTS

Table 13: Assessment of Data Usability For WSP Metrics - Fraser Sockeye. Time series of spawner abundances were assessed using the approach by Brown *et al.* (2020) to determine which WSP metrics are applicable. *Type* identifies whether the available estimates cover the entire CU (Absolute Abundance, *Abs Abd*) or just a subset of the populations (Relative Index, *Rel Idx*). The *Abd* column shows whether the spawner estimates can be used to assess CU abundances (i.e. compared to an absolute benchmark like the COSEWIC threshold for small populations, or to relative benchmarks like 80% of S_{msy}). The *Trend* (*Short- and Long-*) columns show whether the time series has been consistent enough (e.g. survey methods, spatial coverage) to produce meaningful *short-term* and *long-term trend*s. *PercBM* flags whether percentile-based status benchmarks are applicable for the CU, based on the criteria identified by Holt *et al.* (2008).

Species	Area	CU	Type	Abd	Short Trend	Long Trend	Perc BM
Sockeye	Fraser	Anderson_Seton_ES	Abs_Abd	Y	Y	Y	Y
Sockeye	Fraser	Bowron_ES	Abs_Abd	Y	Y	Y	Y
Sockeye	Fraser	Chilko_S_ES	Abs_Abd	Y	Y	Y	Y
Sockeye	Fraser	Chilliwack_ES	Abs_Abd	Y	Y	Y	Y
Sockeye	Fraser	Cultus_L	Abs_Abd	Y	Y	Y	Y
Sockeye	Fraser	Fran_Fras_S	Abs_Abd	Y	Y	Y	Y
Sockeye	Fraser	Harrison_DS_L	Rel_Idx		Y	Y	Y
Sockeye	Fraser	Harrison_R	Abs_Abd	Y	Y	Y	Y
Sockeye	Fraser	Harrison_US_L	Abs_Abd	Y	Y	Y	Y
Sockeye	Fraser	Kamloops_ES	Abs_Abd	Y	Y	Y	Y
Sockeye	Fraser	Lillooet_Harr_L	Abs_Abd	Y	Y	Y	Y
Sockeye	Fraser	Nadina_Francois_ES	Abs_Abd	Y	Y	Y	Y
Sockeye	Fraser	Nahatlatch_ES	Abs_Abd	Y	Y	Y	Y
Sockeye	Fraser	North_Barriere_ES	Abs_Abd	Y	Y	Y	Y
Sockeye	Fraser	Pitt_ES	Abs_Abd	Y	Y	Y	Y
Sockeye	Fraser	Quesnel_S	Abs_Abd	Y	Y	Y	Y
Sockeye	Fraser	Seton_L	Abs_Abd	Y	Y	Y	Y
Sockeye	Fraser	Shuswap_ES	Abs_Abd	Y	Y	Y	Y
Sockeye	Fraser	Shuswap_L	Abs_Abd	Y	Y	Y	Y
Sockeye	Fraser	Takla_Trem_EStu	Abs_Abd	Y	Y	Y	Y
Sockeye	Fraser	Takla_Trem_S_S	Abs_Abd	Y	Y	Y	Y
Sockeye	Fraser	Taseko_ES	Rel_Idx		Y	Y	Y
Sockeye	Fraser	Widgeon_RT	Abs_Abd	Y	Y	Y	Y

Table 14: Metrics and Integrated Status - Fraser Sockeye - Data Up To 2010. This table lists the metric values and resulting statuses from the first integrated status assessment of Fraser River Sockeye (Grant and Pestal 2012). Metrics have been adapted from the original formulation to a format that works more easily with the algorithm calculations (Section 2.2.3). *DataType* identifies whether the available estimates cover the entire CU (Absolute Abundance, Abs Abd) or just a subset of the populations (Relative Index, Rel Idx). *AbsBM* shows the ratio of recent abundance relative to the COSEWIC threshold of 1,000 adults. *LTr* is the *long-term trend*, expressed as the percent change of recent abundance relative to long-term average abundance (e.g. 150 means that recent abundance is 50% larger than long-term; Note that this formulation differs a bit from the other metrics, but is more consistent with the standard WSP benchmark of 50 and 75). *pCh* is the percent change over 3 generations, also called the *percent change* metric. *RelLBM* is the ratio of recent abundance and the lower WSP benchmark for *relative abundance* (S_{gen}). *RelUBM* is the ratio of recent abundance and the upper WSP benchmark for *relative abundance* (80% S_{msy}). *IntStatusRaw* is the integrated status assigned through the expert process, and *IntStatus* is the simplified integrated status used in the algorithm development and testing (i.e. where *Red/Amber* was changed to *Red*, and *Amber/Green* to *Amber*).

ID	CU	Data Type	Abs Met	LTr	pCh	Rel LBM	Rel UBM	Int Status Raw	Int Status
CU-19	Anderson_Seton_ES	Abs_Abd	5.0	175	-39			*Amber*	*Amber*
CU-1	Bowron_ES	Abs_Abd	2.3	28	-88	0.58	0.14	*Red*	*Red*
CU-6	Chilko_S_ES	Abs_Abd	365	123	-74	9.37	1.34	*Green*	*Green*
CU-17	Chilliwack_ES	Abs_Abd	5.8	NA	NA	0.73	0.37	*Red/Amber*	*Red*
CU-8	Cultus_L	Abs_Abd	0.3	9	-69	0.02	0.01	*Red*	*Red*
CU-7	Fran_Fras_S	Abs_Abd	73.3	115	-57	1.74	0.38	*Red/Amber*	*Red*
CU-13	Harrison_DS_L	Rel_Idx	4.4	656	103	NA	NA	*Green*	*Green*
CU-9	Harrison_R	Abs_Abd	115.3	885	3388	12.81	3.6	*Green*	*Green*
CU-11	Harrison_US_L	Abs_Abd	14.4	65	-39	1.6	0.19	*Amber*	*Amber*
CU-2	Kamloops_ES	Abs_Abd	8.5	223	-31	1.41	0.37	*Amber*	*Amber*
CU-12	Lillooet_Harr_L	Abs_Abd	52.9	127	3	4.81	0.69	*Green*	*Green*
CU-3	Nadina_Francois_ES	Abs_Abd	8.8	91	-44	0.52	0.15	*Red*	*Red*
CU-14	Nahatlatch_ES	Abs_Abd	1.7	54	-82			*Red*	*Red*
CU-4	North_Barriere_ES	Abs_Abd	2.7	127	-68	5.22	0.53	*Amber*	*Amber*
CU-5	Pitt_ES	Abs_Abd	22.0	168	-54	3.67	1	*Amber/Green*	*Amber*
CU-22	Quesnel_S	Abs_Abd	56.8	574	-92			*Red/Amber*	*Red*
CU-20	Shuswap_ES	Abs_Abd	19.6	89	-34			*Amber/Green*	*Amber*
CU-23	Shuswap_L	Abs_Abd	35.8	76	46			*Green*	*Green*
CU-18	Takla_Trem_EStu	Abs_Abd	23.1	59	-58			*Red*	*Red*
CU-21	Takla_Trem_S_S	Abs_Abd	44.5	154	-85			*Red/Amber*	*Red*
CU-15	Taseko_ES	Rel_Idx	NA	22	-88			*Red*	*Red*
CU-16	Widgeon_RT	Abs_Abd	0.4	46	736			*Red*	*Red*

Table 15: Metrics and Integrated Status - Fraser Sockeye - Data Up To 2010. This table lists the metric values and resulting statuses from the integrated status re-assessment of Fraser River Sockeye (DFO 2018). Table structure as per Table D.2.

ID	CU	Data Type	Abs BM	LTr	pCh	Rel LBM	Rel UBM	Int Status Raw	Int Status
CU-19	Anderson_Seton_ES	Abs_Abd	18.0	684	430	4.93	0.80	*Amber/Green*	*Amber*
CU-1	Bowron_ES	Abs_Abd	1.7	35	-37	0.33	0.09	*Red*	*Red*
CU-6	Chilko_S_ES	Abs_Abd	599	246	173	8.28	1.58	*Green*	*Green*
CU-17	Chilliwack_ES	Abs_Abd	12.6	158	102	1.58	0.79	*Amber/Green*	*Amber*
CU-8	Cultus_L	Abs_Abd	0.4	14	3	0.03	0.01	*Red*	*Red*
CU-7	Fran_Fras_S	Abs_Abd	135	136	-21	5.57	1.10	*Amber/Green*	*Amber*
CU-13	Harrison_DS_L	Rel_Idx	4.2	445	-75			*Amber/Green*	*Amber*
CU-9	Harrison_R	Abs_Abd	165	1924	2007	4.25	1.35	*Green*	*Green*
CU-11	Harrison_US_L	Abs_Abd	5.4	45	-82	0.50	0.06	*Red*	*Red*
CU-2	Kamloops_ES	Abs_Abd	10.7	169	-40	2.16	0.60	*Amber*	*Amber*
CU-12	Lillooet_Harr_L	Abs_Abd	28.8	81	-72	2.03	0.37	*Amber*	*Amber*
CU-3	Nadina_Francois_ES	Abs_Abd	26.1	151	123	1.20	0.38	*Amber/Green*	*Amber*
CU-14	Nahatlatch_ES	Abs_Abd	2.3	100	16			*Amber*	*Amber*
CU-4	North_Barriere_ES	Abs_Abd	2.8	121	-52	4.38	0.55	*Amber*	*Amber*
CU-5	Pitt_ES	Abs_Abd	47.7	234	-44	7.64	2.35	*Green*	*Green*
CU-22	Quesnel_S	Abs_Abd	33.2	221	-95	0.96	0.13	*Red/Amber*	*Red*
CU-10	Seton_L	Abs_Abd	0.6	31	-91	0.27	0.04	*Red*	*Red*
CU-20	Shuswap_ES	Abs_Abd	9.5	190	145	2.90	0.75	*Amber*	*Amber*
CU-23	Shuswap_L	Abs_Abd	12.5	46	-85	6.80	1.18	*Amber/Green*	*Amber*
CU-18	Takla_Trem_EStu	Abs_Abd	26.8	43	-44	0.77	0.20	*Red*	*Red*
CU-21	Takla_Trem_S_S	Abs_Abd	44.5	112	-63	1.16	0.21	*Red/Amber*	*Red*
CU-15	Taseko_ES	Rel_Idx	NA	25	-81			*Red*	*Red*
CU-16	Widgeon_RT	Abs_Abd	0.4	178	1158			*Red*	*Red*

C.2 SOUTHERN BC CHINOOK DATA USABILITY, METRIC VALUE, AND INTEGRATED STATUS ASSESSMENT RESULTS

Table 16: Assessment of Data Usability For WSP Metrics – Fraser Chinook. Time series of spawner abundances were assessed by Brown *et al.* (2020) to determine which WSP metrics are applicable. *Type* identifies whether the available estimates cover the entire CU (Absolute Abundance, *Abs Abd*) or just a subset of the populations (Relative Index, *Rel Idx*). The *Abd* column shows whether the spawner estimates can be used to assess CU abundances (i.e. compared to an absolute benchmark like the COSEWIC threshold for small populations, or to relative benchmarks like 80% of S_{msy}). The *Trend* columns show whether the time series has been consistent enough (e.g. survey methods, spatial coverage) to produce meaningful short-term and *long-term trends*. *PercBM* flags whether percentile-based status benchmarks are applicable for the CU, based on the criteria identified by Holt *et al.* (2008).

Species	Area	CU	Type	Abd	Short Trend	Long Trend	Perc BM
Chinook	Fraser	Lower Fraser River_FA_0.3	Abs_Abd	Y	Y	Y	
Chinook	Fraser	Lower Fraser River_SP_1.3	Rel_Idx		Y	Y	
Chinook	Fraser	Lower Fraser River_SU_1.3	Rel_Idx			Y	
Chinook	Fraser	Lower Fraser River-Upper Pitt_SU_1.3	Rel_Idx			Y	
Chinook	Fraser	Lower Thompson_SP_1.2	Rel_Idx		Y	Y	
Chinook	Fraser	Maria Slough_SU_0.3	Rel_Idx			Y	
Chinook	Fraser	Middle Fraser River_SP_1.3	Rel_Idx		Y	Y	
Chinook	Fraser	Middle Fraser River_SU_1.3	Rel_Idx			Y	
Chinook	Fraser	Middle Fraser River-Portage_FA_1.3	Rel_Idx		Y	Y	
Chinook	Fraser	Middle Fraser-Fraser Canyon_SP_1.3	Rel_Idx			Y	
Chinook	Fraser	North Thompson_SP_1.3	Rel_Idx			Y	
Chinook	Fraser	North Thompson_SU_1.3	Rel_Idx		Y	Y	
Chinook	Fraser	Shuswap River_SU_0.3	Abs_Abd	Y	Y	Y	
Chinook	Fraser	South Thompson_SU_0.3	Rel_Idx		Y	Y	
Chinook	Fraser	South Thompson_SU_1.3	Rel_Idx		Y	Y	
Chinook	Fraser	South Thompson-Bessette Creek_SU_1.2	Rel_Idx			Y	
Chinook	Fraser	Upper Adams River_SU_x.x	Rel_Idx			Y	
Chinook	Fraser	Upper Fraser River_SP_1.3	Rel_Idx		Y	Y	

Table 17: Assessment of Data Usability For WSP Metrics – Other Chinook. Time series of spawner abundances were assessed by Brown *et al.* (2020) to determine which WSP metrics are applicable. *Type* identifies whether the available estimates cover the entire CU (Absolute Abundance, *Abs Abd*) or just a subset of the populations (Relative Index, *Rel Idx*). The *Abd* column shows whether the spawner estimates can be used to assess CU abundances (i.e. compared to an absolute benchmark like the COSEWIC threshold for small populations, or to relative benchmarks like 80% of S_{msy}). The *Trend* columns show whether the time series has been consistent enough (e.g. survey methods, spatial coverage) to produce meaningful short-term and *long-term trends*. *PercBM* flags whether percentile-based status benchmarks are applicable for the CU, based on the criteria identified by Holt *et al.* (2008).

Species	Area	CU	Type	Abd	Short Trend	Long Trend	Perc BM
Chinook	Columbia	Okanagan_1.x	Abs_Abd	Y	Y	Y	
Chinook	Inner South Coast	Boundary Bay_FA_0.3	Rel_Idx		Y	Y	
Chinook	Inner South Coast	East Vancouver Island - Georgia Strait Summer 0.3	Rel_Idx				
Chinook	Inner South Coast	East Vancouver Island-Cowichan & Koksilah_FA_0.x	Abs_Abd	Y	Y	Y	
Chinook	Inner South Coast	East Vancouver Island-Goldstream_FA_0.x	Abs_Abd			Y	
Chinook	Inner South Coast	East Vancouver Island-Nanaimo & Chemainus_FA_0.x	Rel_Idx			Y	
Chinook	Inner South Coast	East Vancouver Island-Nanaimo_SP_1.x	Rel_Idx			Y	
Chinook	Inner South Coast	East Vancouver Island-North_FA_0.x	Rel_Idx		Y	Y	
Chinook	Inner South Coast	East Vancouver Island-Qualicum & Puntledge_FA_0.x	Rel_Idx			Y	
Chinook	Inner South Coast	Homathko_SU_x.x	Rel_Idx				
Chinook	Inner South Coast	Klinaklini_SU_1.3	Rel_Idx				
Chinook	Inner South Coast	Southern Mainland-Georgia Strait_FA_0.x	Rel_Idx			Y	
Chinook	Inner South Coast	Southern Mainland-Southern Fjords_FA_0.x	Rel_Idx			Y	
Chinook	WCVI	West Vancouver Island-Nootka & Kyuquot_FA_0.x	Rel_Idx		Y	Y	
Chinook	WCVI	West Vancouver Island-North_FA_0.x	Rel_Idx			Y	
Chinook	WCVI	West Vancouver Island-South_FA_0.x	Rel_Idx		Y	Y	

Table 18: Metrics and Integrated Status – Southern BC Chinook - Data Up To 2012. This table lists the metric values and resulting statuses from the integrated status assessment of Southern BC Chinook (DFO 2016). Table structure as per Table D.2.2*.*

ID	CU	Data Type	Abs BM	LTr	pCh	Rel LBM	Rel UBM	Int Status Raw	Int Status
CK-29	East Vancouver Island-North_FA_0.X	Rel_Idx		110	-2			*Red*	*Red*
CK-03	Lower Fraser River_FA_0.3	Abs_Abd	79.7	86	-51	1.86	1.12	*Green*	*Green*
CK-17	Lower Thompson_SP_1.2	Rel_Idx		53	-79			*Red*	*Red*
CK-10	Middle Fraser River_SP_1.3	Rel_Idx		61	-68			*Red*	*Red*
CK-11	Middle Fraser River_SU_1.3	Rel_Idx		81	-48			*Amber*	*Amber*
CK-09	Middle Fraser River-Portage_FA_1.3	Rel_Idx		77	-67			*Red*	*Red*
CK-18	North Thompson_SP_1.3	Rel_Idx		54	-87			*Red*	*Red*
CK-19	North Thompson_SU_1.3	Rel_Idx		59	-75			*Red*	*Red*
CK-01	Okanagan_1.x	Rel_Idx		60	-33			*Red*	*Red*
CK-13	South Thompson_SU_0.3	Rel_Idx		128	40			*Green*	*Green*
CK-14	South Thompson_SU_1.3	Rel_Idx		87	-57			*Red/ Amber*	*Red*
CK-16	South Thompson-Bessette Creek_SU_1.2	Rel_Idx		25	-95			*Red*	*Red*
CK-12	Upper Fraser River_SP_1.3	Rel_Idx		65	-62			*Red*	*Red*
CK-32	West Vancouver Island-Nootka & Kyuquot_FA_0.X	Rel_Idx		75	-32			*Red*	*Red*
CK-31	West Vancouver Island-South_FA_0.X	Rel_Idx		73	-18			*Red*	*Red*

C.3 INTERIOR FRASER COHO DATA USABILITY, METRIC VALUE, AND INTEGRATED STATUS ASSESSMENT RESULTS

Table 19: Assessment of Data Usability For WSP Metrics – Interior Fraser Coho. Time series of spawner abundances were assessed using the approach by Brown *et al.* (2020) to determine which WSP metrics are applicable. *Type* identifies whether the available estimates cover the entire CU (Absolute Abundance, *Abs Abd*) or just a subset of the populations (Relative Index, *Rel Idx*). The *Abd* column shows whether the spawner estimates can be used to assess CU abundances (i.e. compared to an absolute benchmark like the COSEWIC threshold for small populations, or to relative benchmarks like 80% of S_{msy}). The *Trend* columns show whether the time series has been consistent enough (e.g. survey methods, spatial coverage) to produce meaningful short-term and *long-term trends*. *PercBM* flags whether percentile-based status benchmarks are applicable for the CU, based on the criteria identified Holt *et al.* (2008).

Species	Area	CU	Type	Abd	Short Trend	Long Trend	Perc BM
Coho	Fraser	Middle Fraser	Abs_Abd	Y	Y	Y	
Coho	Fraser	Fraser Canyon	Abs_Abd	Y	Y	Y	
Coho	Fraser	Lower Thompson	Abs_Abd	Y	Y	Y	
Coho	Fraser	North Thompson	Abs_Abd	Y	Y	Y	
Coho	Fraser	South Thompson	Abs_Abd	Y	Y	Y	

Table 20: Metrics and Integrated Status – Interior Fraser Coho - Data Up To 2013. This table lists the metric values and resulting statuses from the integrated status assessment of Interior Fraser Coho. (DFO 2013). Table structure as per Table D.2.2.

ID	CU	Data Type	Abs BM	LTr	pCh	Rel LBM	Rel UBM	Int Status Raw	Int Status
CO-02	Fraser Canyon	Abs_Abd	4.5	135	-13	6.01	2.85	*Amber*	*Amber*
CO-03	Lower Thompson	Abs_Abd	9.2	232	452	6.54	3.01	*Amber/ Green*	*Amber*
CO-01	Middle Fraser	Abs_Abd	6.9	193	62	4.37	2.48	*Amber*	*Amber*
CO-04	North Thompson	Abs_Abd	13.7	98	83	5.37	2.58	*Amber/ Green*	*Amber*
CO-05	South Thompson	Abs_Abd	8.1	121	40	3.43	1.83	*Amber*	*Amber*

APPENDIX D: CANDIDATE ALGORITHMS AND SPLITTING RULES

D.1 MINIMALIST

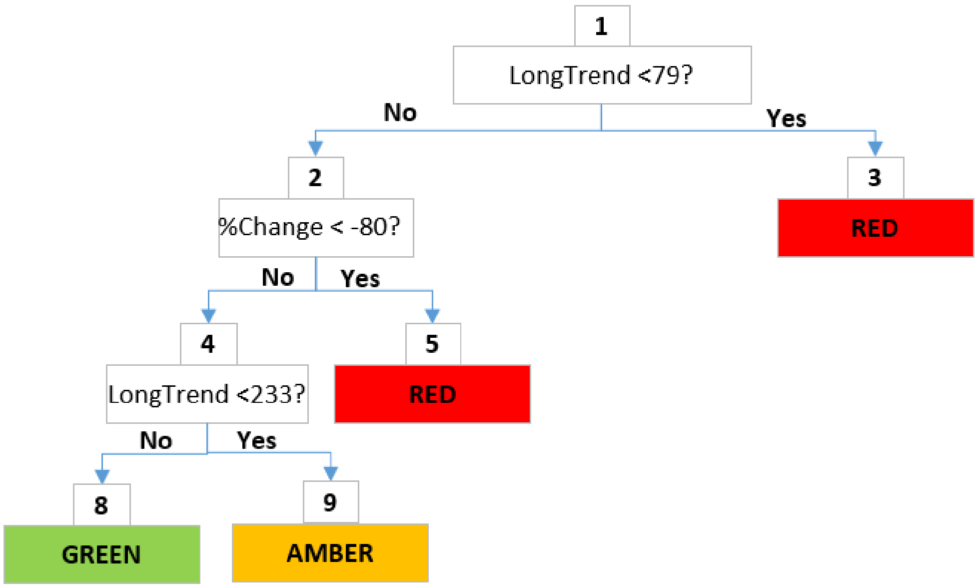


Figure 23. Classification Tree – *Minimalist.*

Table 21: Classification Rules - *Minimalist.*

Node	Status	Rule
Node3	*Red*	LongTrend < 79
Node5	*Red*	LongTrend >= 79, then PercChange < -80
Node8	*Green*	LongTrend >= 79, then PercChange >= -80, then LongTrend >= 233
Node9	*Amber*	LongTrend >= 79, then PercChange >= -80, then LongTrend < 233

D.2 FANCY PANTS

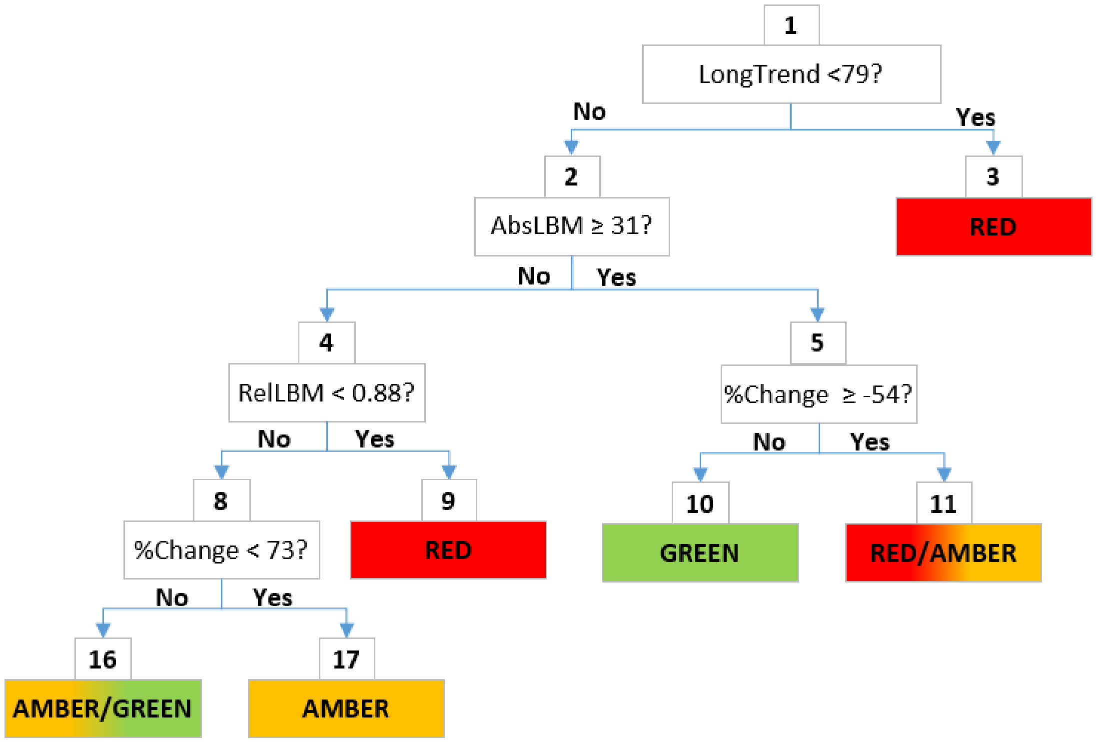


Figure 24. Classification Tree – Fancy Pants.

Table 22: Classification Rules – *Fancy Pants.*

Node	Status	Rule
Node3	*Red*	LongTrend < 79
Node10	*Green*	LongTrend >= 79, then AbsLBM >= 31, then PercChange >= -54
Node11	*Red/Amber*	LongTrend >= 79, then AbsLBM >= 31, then PercChange < -54
Node9	*Red*	LongTrend >= 79, then AbsLBM < 31, then RelLBM < 0.88
Node16	*Amber*	LongTrend >= 79, then AbsLBM < 31, then RelLBM < 0.88, then PercChange < 73
Node17	*Amber/Green*	LongTrend >= 79, then AbsLBM < 31, then RelLBM < 0.88, then PercChange >= 73

D.3 CATEGORICAL REALIST

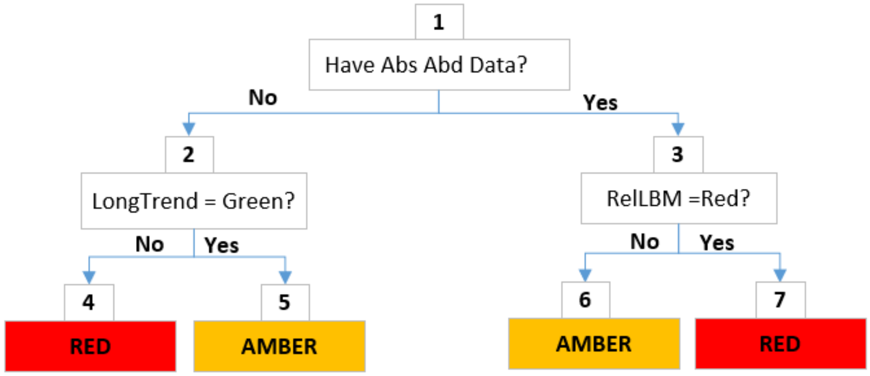


Figure 25. Classification Tree – Categorical Realist.

Table 23: Classification Rules – *Categorical Realist.*

Node	Status	Rule
Node4	*Amber*	DataType is Rel_idx, then LongTrend is *Amber*
Node5	*Red*	DataType is Rel_idx, then LongTrend is *Red* or *Amber*
Node6	*Amber*	DataType is Abs_Abd, then RelLBM is *Amber* or *Amber*
Node7	*Red*	DataType is Abs_Abd, then RelLBM is *Red*

D.4 SIMPLY *RED*

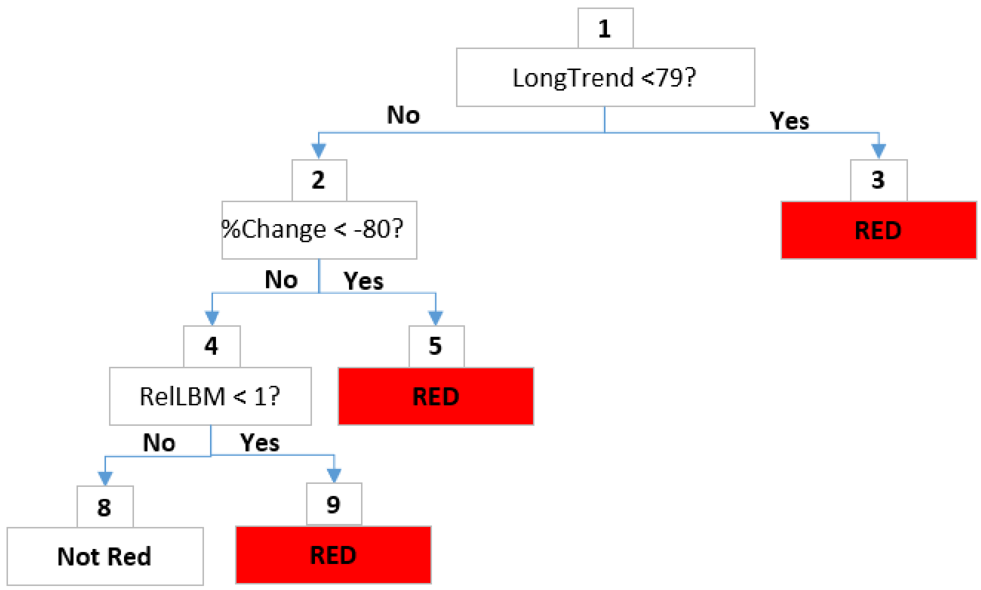


Figure 26. Classification Tree – Simply Red.

Table 24: Classification Rules – *Simply Red.*

Node	Status	Rule
Node3	*Red*	LongTrend < 79
Node5	*Red*	LongTrend >= 79, then PercChange < -70
Node8	*Not Red*	LongTrend >= 79, then PercChange >= -70, then RelLBM >= 1
Node9	*Red*	LongTrend >= 79, then PercChange >= -70, then RelLBM < 1

D.5 LEARNING TREE 1

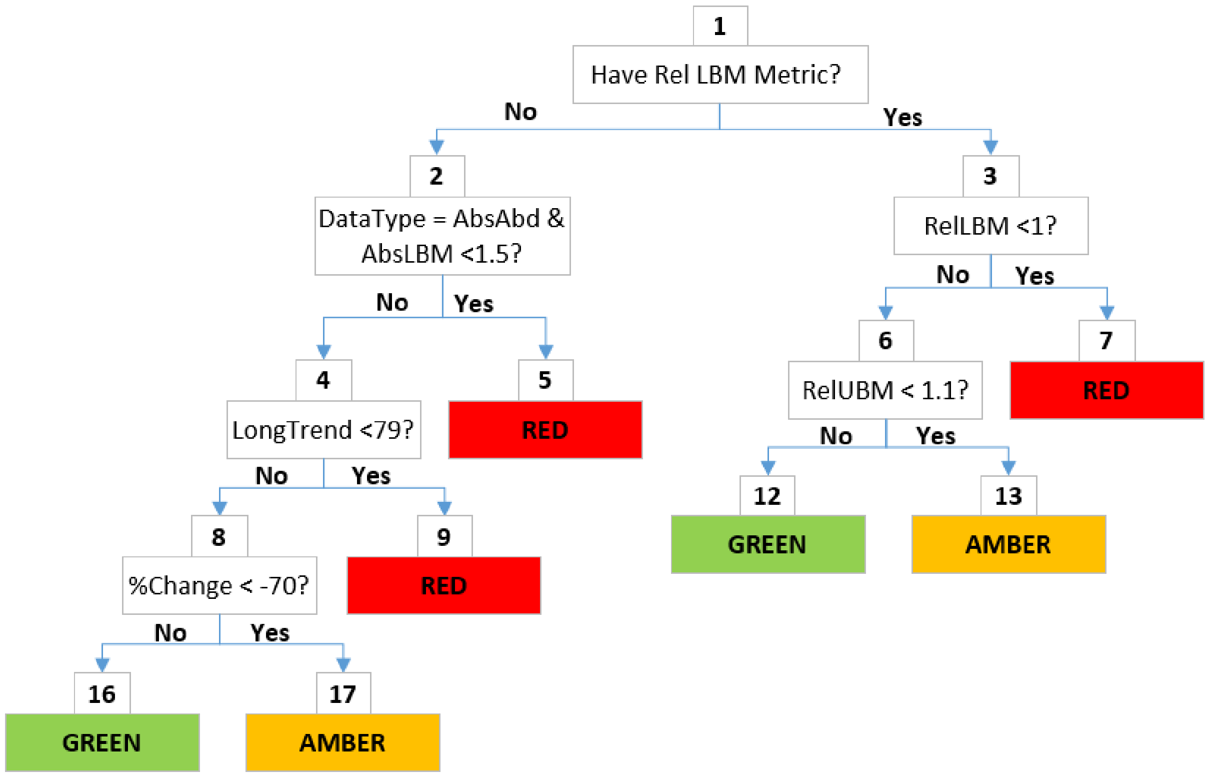


Figure 27. Classification Tree – Learning Tree 1.

Table 25: Classification Rules – *Learning Tree 1.*

Node	Status	Rule
Node7	*Red*	Have RelLBM, then RelLBM < 1
Node12	*Green*	Have RelLBM, then RelLBM >= 1, then RelUBM >= 1.1
Node13	*Amber*	Have RelLBM, then RelLBM >= 1, then RelUBM < 1.1
Node5	*Red*	Don't have RelLBM, then Data Type is Abs_Abd AND AbsLBM < 1.5
Node9	*Red*	Don't have RelLBM, then Data Type is Rel_idx OR AbsLBM >= 1.5, then LongTrend < 79
Node16	*Green*	Don't have RelLBM, then Data Type is Rel_idx OR AbsLBM >= 1.5, then LongTrend >= 79, then PercChange >= -70
Node17	*Amber*	Don't have RelLBM, then Data Type is Rel_idx OR AbsLBM >= 1.5, then LongTrend >= 79, then PercChange < -70

D.6 LEARNING TREE 2

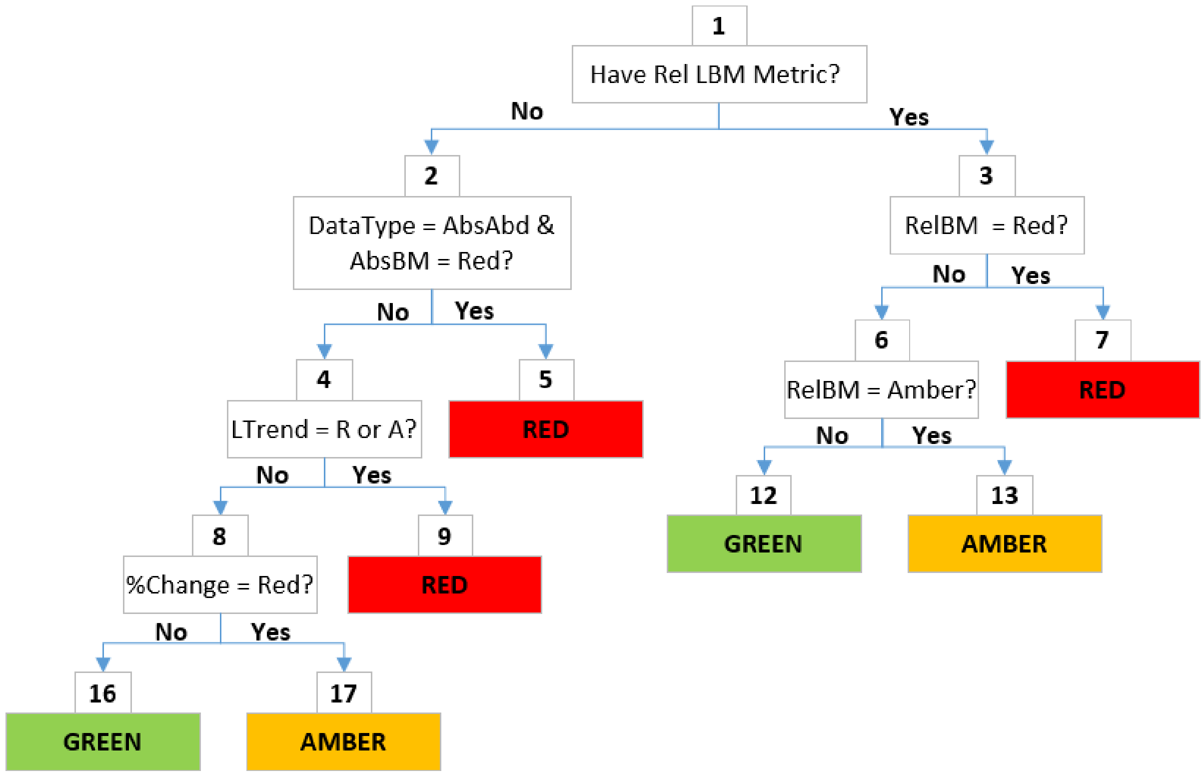


Figure 28. Classification Tree – *Learning Tree 2.*

Table 26: Classification Rules – *Learning Tree 2.*

Node	**Status**	**Rule**
Node7	*Red*	Have RelBM, then RelBM is *Red*
Node12	*Green*	Have RelBM, then RelBM is *Amber* or *Amber*, then RelBM is *Amber*
Node13	*Amber*	Have RelBM, then RelBM is *Amber* or *Amber*, then RelBM is *Amber*
Node5	*Red*	Don't have RelBM, then Data Type is AbsAbd and AbsBM is *Red*
Node9	*Red*	Don't have RelBM, then Data Type is Rel_idx OR AbsBM is *Amber* or *Amber*, then LongTrend is *Red* or *Amber*
Node16	*Green*	Don't have RelBM, then Data Type is Rel_idx OR AbsBM is *Amber* or *Amber*, then LongTrend is *Amber*, then PercChange is *Amber* or *Amber*
Node17	*Amber*	Don't have RelBM, then Data Type is Rel_idx OR AbsBM is *Amber* or *Amber*, then LongTrend is *Amber*, then PercChange is *Red*

D.7 LEARNING TREE 3

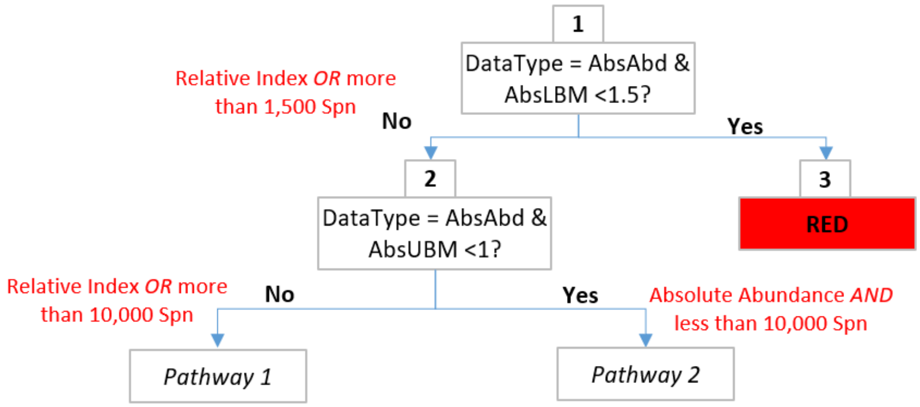


Figure 29. Classification Tree – Learning Tree 3: Initial Steps.

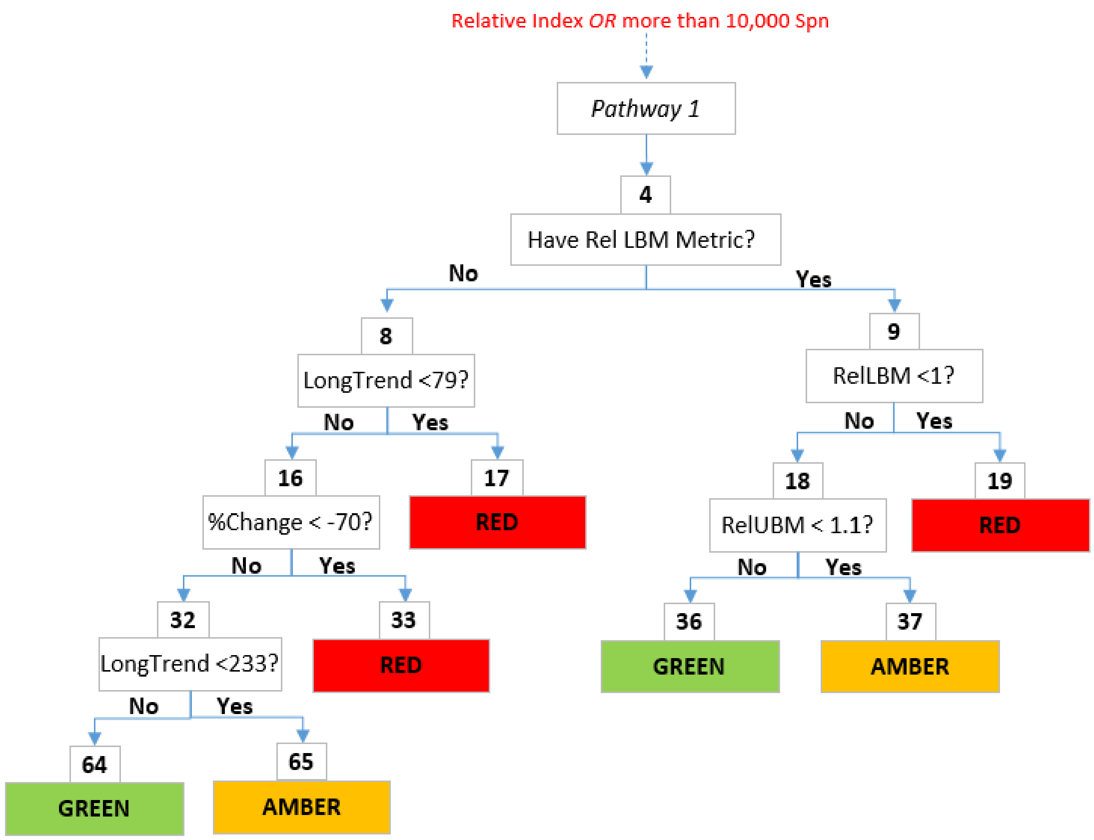


Figure 30. Classification Tree – Learning Tree 3: Pathway 1.

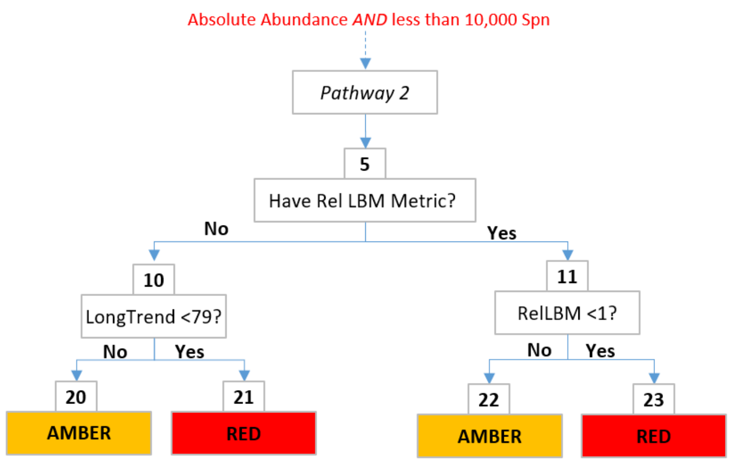


Figure 31. Classification Tree – Learning Tree 3: Pathway 2.

Table 27: Classification Rules – *Learning Tree 3.*

Node	Status	Rule
Node3	*Red*	Data Type is AbsAbd AND AbsLBM < 1.5
Node17	*Red*	Data Type is RelIdx OR AbsLBM >= 1.5, then Data Type is RelIdx OR AbsUBM >= 1, then Don't have RelLBM, then LongTrend < 79
Node19	*Red*	Data Type is RelIdx OR AbsLBM >= 1.5, then Data Type is RelIdx OR AbsUBM >= 1, then Have RelLBM, then RelLBM < 1
Node20	*Amber*	Data Type is RelIdx OR AbsLBM >= 1.5, then Data Type is AbsAbd AND AbsUBM < 1,then Don't have RelLBM, then LongTrend >= 79
Node21	*Red*	Data Type is RelIdx OR AbsLBM >= 1.5, then Data Type is AbsAbd AND AbsUBM < 1,then Don't have RelLBM, then LongTrend < 79
Node22	*Amber*	Data Type is RelIdx OR AbsLBM >= 1.5, then Data Type is AbsAbd AND AbsUBM < 1,then Have RelLBM, then RelLBM >= 1
Node23	*Red*	Data Type is RelIdx OR AbsLBM >= 1.5, then Data Type is AbsAbd AND AbsUBM < 1,then Have RelLBM, then RelLBM < 1
Node33	*Red*	Data Type is RelIdx OR AbsLBM >= 1.5, then Data Type is RelIdx OR AbsUBM >= 1, then Don't have RelLBM, then LongTrend >= 79, then PercChange < -70
Node36	*Green*	Data Type is RelIdx OR AbsLBM >= 1.5, then Data Type is RelIdx OR AbsUBM>=1, then have RelLBM,then RelLBM>=1, then RelUBM>=1.1
Node37	*Amber*	Data Type is RelIdx OR AbsLBM >= 1.5, then Data Type is RelIdx OR AbsUBM>= 1,then Have RelLBM, then RelLBM >= 1,then RelUBM<1.1
Node64	*Green*	Data Type is RelIdx OR AbsLBM >= 1.5, then Data Type is RelIdx OR AbsUBM >= 1, then Don't have RelLBM, then LongTrend >= 79, then PercChange >= -70, then LongTrend >= 233
Node65	*Amber*	Data Type is RelIdx OR AbsLBM >= 1.5, then Data Type is RelIdx OR AbsUBM >= 1, then Don't have RelLBM, then LongTrend >= 79, then PercChange >= -70, then LongTrend < 233

APPENDIX E: PERFORMANCE OF INDIVIDUAL ALGORITHMS WITH THE LEARNING DATA SET

Results in this section refer to the performance summaries for all cases (

Table 7) and by species (

Table 8 to Table 10), as well as the detailed error diagnostics for each algorithm.

All fitted CART algorithms and the Simply *Red* constructed algorithm do not meet *Criteria 5 and 6*, since they rely on the analyses, rather than WSP status integration logic. Only the *Learning Trees* meet these two criteria.

Note that throughout this Appendix we present three different performance metrics that are expressed as percentage values for comparison:

- Correct cases, expressed as a % of total cases. We include this at the beginning of each section as an overall summary for easy comparison between algorithms, but it conflates the two distinct considerations captured in criteria 1 and 3.
- Completed cases, where the algorithm could assign a rapid status, expressed as % of total cases. This is used in *Criterion 3*.
- Correct cases, expressed as % of completed cases. This is used in *Criterion 1*.

E.1 FITTED ALGORITHM: MINIMALIST

Of the 65 total cases in the *learning data set*, the *Minimalist* could assign a rapid status to 64 of them (99%) and assigned the correct status to 49 of them on the 3-status scale (75%).

Criterion 1: the *Minimalist* is relatively accurate, assigning 49/64 (77%) of the completed cases correctly on its 3 status scale: *Red*, *Amber*, and *Green*. Positive prediction errors, where algorithm statuses were better than the WSP integrated statuses, all deviated by only one status zone.

Criterion 2: Errors are roughly balanced between predicting better (8/64: 13%) versus poorer (7/64: 11%) rapid statuses than the associated WSP integrated statuses, so the algorithm does not err on the side of being precautionary with the test cases.

Criterion 3: It uses only trend-based metrics, so it is applicable across all data types available for Pacific salmon CUs. The *Minimalist* assigns rapid statuses for almost all the *learning data set* CUs (64/65: 98%).

Criterion 4: It predicts the three main status zones: *Red*, *Amber* and *Green*, so it meets this criterion.

Criterion 5: Thresholds estimated using CART, so does not explicitly meet this criterion.

Criterion 6: Criteria and their sequences estimated using CART, so does not explicitly meet this criterion.

E.2 FITTED ALGORITHM: FANCY PANTS

Of the 65 total cases in the *learning data set*, *Fancy Pants* could assign a rapid status to 54 of them (83%) and assigned the correct status to 47 of them on the 5-status scale (72%).

Criterion 1: The *Fancy Pants* algorithm is relatively accurate, assigning 47/54 (87%) completed rapid statuses correctly on its 5 status scale: *Red*, *Red/Amber*, *Amber*, *Amber/Green* and *Green.* Accuracy increases slightly on the 3 status scale to 49/54 (91%). Most of the 7 incorrectly assigned statuses were off by 1 status zone (50/54: 93%).

Criterion 2: There were more errors that predicted better 12/54 (22%) than worse 4/54 (7%) rapid statuses compared to the WSP integrated status; so this algorithm is not precautionary and does not meet this criterion.

Criterion 3: It assigns rapid statuses to a high proportion (54/65: 83%) of the *learning data set* CUs. This high accuracy is a product of the data we have available for testing, which is heavily weighted towards data-rich Fraser sockeye CUs. This algorithm highly depends on the availability of the *absolute abundance* metric to assess status. This condition limits the applicability of *Fancy Pants*, since absolute abundance data, required to estimate the *absolute abundance* metric, are not available for most CUs in the Pacific Region. For Southern BC Chinook, for example, absolute abundance data are available for only one CU. Apart from this one CU, the *Fancy Pants* algorithm could only produce statuses for Southern BC Chinook CUs that had a *Long Term Trend* metric value that was less than 79%. The remaining Southern BC Chinook CUs could not be assigned statuses. Therefore, this algorithm has limited applicability more broadly.

Criterion 4: It predicts the all status zones: *Red*, *Red/Amber*, *Amber*, *Amber/Green* and *Green*, so it meets this criterion.

Criterion 5: Thresholds estimated using CART, so does not explicitly meet this criterion.

Criterion 6: Criteria *and* their sequences estimated using CART, so does not explicitly meet this criterion.

E.3 FITTED ALGORITHM: CATEGORICAL REALIST

Of the 65 total cases in the *learning data set*, *Categorical Realist* could assign a rapid status to 55 of them (85%) and assigned the correct status to 41 of them on the 3-status scale (63%).

Criterion 1: The *Categorical Realist* algorithm is relatively accurate, assigning 41/55 (75%) rapid statuses correctly on its 3 status scale.

Criterion 2: Only 5/55 (9%) errors predicted better statuses, and these were limited to one status zone higher than the WSP integrated statuses. This is by design, since only two status zones are assigned by the algorithm. More predicted statuses were worse than the WSP integrated statuses (9/55: 16%), which is precautionary, meeting this criterion.

Criterion 3: It assigns rapid statuses to a high proportion (55/65: 85%) of the *learning data set* CUs. Therefore this algorithm has broad applicability across Cus.

Criterion 4: Categorical Realist only predicts *Amber* and *Red* status zones, so does not meet this criterion.

Criterion 5: Thresholds estimated using CART, so does not explicitly meet this criterion.

Criterion 6: Criteria and their sequences estimated using CART, so does not explicitly meet this criterion.

E.4 CONSTRUCTED ALGORITHM: SIMPLY *RED*

Of the 65 total cases in the *learning data set*, *Categorical Realist* could assign a rapid status to 55 of them (85%) and assigned the correct status to 47 of them on the 2-status scale (72%).

Criterion 1: Simply *Red* is relatively accurate, assigning 47/55 (85%) rapid status correctly on its 2 status scale: *Red* and *NotRed*. When results from this algorithm are compared on the 3 status scale, the number correct drops by half: 26/55 (47%).

Criterion 2: There were more errors on the 2-status scale that predict poorer status than errors that predict better status than the integrated assessment (2 predicted better, out of 8 errors). Even on the 3 status scale most of the errors on are only +1 or -1, indicating that *Simply Red* assigned a *NotRed* status to cases with a WSP integrated status of either *Amber* or *Green*, which is essentially correct but is scored as an error on the 3-status scale..

Criterion 3: It assigns rapid statuses for 55/65 (85%) of the *learning data set* CUs*.* This algorithm is limited by its reliance on the *relative abundance* metric to get to a *NotRed* status. This reliance means that the algorithm is not as applicable as others across the range of data types present.

Criterion 4: This algorithm assigns only *Red* and *NotRed* statuses, therefore, does not meet this criterion.

Criterion 5: Built from components of fitted trees. Thresholds estimated using CART, so does not explicitly meet this criterion.

Criterion 6: Built from components of fitted trees. Criteria and their sequences estimated using CART, so does not explicitly meet this criterion.

E.5 CONSTRUCTED ALGORITHMS: LEARNING TREES 1 & 2

Differences in rapid statuses assigned by *Learning Tree 1* versus *Learning Tree* 2 are generally the result of the precautionary nature of the metric thresholds in *Learning Tree* 1. Specifically, this rapid status algorithm’s thresholds are more biologically conservative than the WSP benchmarks used to delineate metric status zones, with the exception of the *percent change* threshold. The *percent change* tree node in *Learning Tree* 1 has a less biologically conservative threshold than the associated WSP benchmark and is responsible for the poorer performance of this algorithm on Criterion 2 compared to *Learning Tree* 2.

<u>Learning Tree 1</u>

Of the 65 total cases in the *learning data set*, *Learning Tree 1* could assign a rapid status to 65 of them (100%) and assigned the correct status to 46 of them on the 3-status scale (71%).

Criterion 1: Learning Tree 1 correctly assigns rapid statuses for 46/65 (71%) cases.

Criterion 2: Of the assessed cases, 17/65 (26%) were assigned a better rapid status than the WSP integrated status. Hence, this algorithm produces less precautionary status results, and therefore, does not perform well on Criterion 2.

Criterion 3: This algorithm assigns rapid statuses for 65/65 (100%) cases in the *learning data set* CUs.

Criterion 4: It predicts the three main status zones: *Red*, *Amber* and *Green*; meets this.

Criterion 5: Constructed based on review of thresholds from the fitted trees, to ensure that this criterion is met.

Criterion 6: Constructed based on review of criteria and sequence from the fitted trees in combination with the status narratives from the expert workshops to ensure that this criterion is met.

Learning Tree 2

Of the 65 total cases in the *learning data set*, *Learning Tree 1* could assign a rapid status to 65 of them (100%) and assigned the correct status to 48 of them on the 3-status scale (74%).

Criterion 1: Learning Tree 2 correctly assigns rapid statuses for 48/65 (74%) cases, improving slightly on the *Learning Tree* 1 in terms of accuracy.

Criterion 2: Similar to *Learning Tree* 1, this algorithm assigned a better statuses to 16/65 (25%) cases, compared to their WSP integrated statuses. Hence, this algorithm produces less precautionary status results, and therefore, does not perform well on Criterion 2.

Criterion 3: This algorithm assign rapid statuses for 65/65 (100%) cases in the *learning data set* CUs.

Criterion 4: It predicts the three main status zones: *Red*, *Amber* and *Green*, so it meets this criterion.

Criterion 5: Constructed based on review of thresholds from the fitted trees, to ensure that this criterion is met.

Criterion 6: Constructed based on review of criteria and sequence from the fitted trees in combination with tatus narratives from expert workshops to ensure that this criterion is met.

E.6 CONSTRUCTED ALGORITHM: LEARNING TREE 3

Of the 65 total cases in the *learning data set*, *Learning Tree 1* could assign a rapid status to 65 of them (100%) and assigned the correct status to 54 of them on the 3-status scale (83%).

Criterion 1: Learning Tree 3 correctly assigns rapid statuses for 54/65 (83%) cases. This is the highest overall number of accurate rapid status assignments across all candidate algorithms. Therefore, *Learning Tree 3* improves upon *Learning Tree*s 1 and 2 in terms of accuracy.

Criterion 2: Of the incorrect statuses, only 7/65 (11%) cases were better than the WSP integrated statuses. Hence, this tree improves upon *Learning Tree*s 1 and 2 in terms of adherence to the precautionary approach.

Criterion 3: It assigns rapid statuses for 65/65 (100%) cases in the *learning data set* CUs.

Criterion 4: It predicts the three main status zones: *Red*, *Amber* and *Green*, so it meets this criterion.

Criterion 5: Constructed based on review of thresholds from the fitted trees, to ensure that this criterion is met.

Criterion 6: Constructed based on review of criteria and sequence from the fitted trees in combination with the status narratives from the expert workshops to ensure that this criterion is met.

Learning Tree 3 also was the most robust of the three *Learning Tree* variations in the *relative abundance* metric sensitivity test (Section 3.2.5).

APPENDIX F: DETAILED RESULTS FOR LEARNING DATA SET

Algorithm performance is assessed by comparing the WSP rapid status to WSP integrated statuses, analogous to evaluating regression fits based on magnitude and pattern of residuals.

Table 7 to Table 10 compare error distributions across candidate algorithms. This appendix includes more detailed error diagnostics for each algorithm, summarized in three different ways:

- *Confusion Matrix*: A cross-tabulation of integrated statuses against rapid statuses assigned by the algorithm, for all cases in the Learning Set.
- *Error Type Frequency*: Number and percentage of different error types (e.g. predicted status better than integrated status, predicted status worse than integrated status). Values are shown for all cases combined, and by species.
- *Error Score Frequency*: Similar to the Error Type Frequency, but based on converting statuses to numeric equivalents (1 = *Red*, 5 = *Green*). Values are shown for all cases combined, and by species.

Note that the error diagnostics in this Appendix use the most appropriate status scale for each algorithm:

- *3-status scale* (*Red*, *Amber*, *Green*): *Minimalist*, *Categorical Realist, Learning Tree 1-3*
- *5-status scale* (*Red*, *Red/Amber*, *Amber*, *Amber/Green*, *Green*): *Fancy Pants*
- *2 status scale* (*Red*, *NotRed*): *Simply Red*

F.1 ERROR DIAGNOSTICS - MINIMALIST

Table 28: Minimalist – Confusion Matrix. Cross-tabulation of integrated statuses against WSP rapid statuses (Predicted). Numbers shown are the cases for each combination of WSP integrated status and rapid statuses. Yellow shaded cells show cases where the rapid status is considered correct on the appropriate error scale for this algorithm.

Integrated Status

Predicted	*Red*	*Amber*	*Green*
Red	25	2	1
Amber	6	19	4
Amber	0	2	5
None	1	0	0

Table 29: Minimalist – Error Type Frequency. Table shows four types of error. *NA* errors are cases where a rapid status could not be assigned with the algorithm. *None* is the number of cases without error, where the rapid status matches the integrated status. *PredBetter* is the number of cases where the rapid status is a better status than the integrated status. *PredWorse* is the number of cases where the rapid status is a poorer status than the integrated status*. PercCorrect*, *PercBetter*, and *PercWorse* are the percent of each error type (i.e. number of cases relative to total number of cases excluding NA cases).

ErrorType	All	Chinook	Coho	Sockeye
NA	1	NA	NA	1
None	49	11	5	33
PredBetter	8	2	NA	6
PredWorse	7	2	NA	5
Total(excl.NA)	64	15	5	44
PercCorrect	77	73	100	75
PercWorse	11	13	NA	11
PercBetter	12	13	NA	14

Table 30: Minimalist – Error Scores Frequency. Table shows the number of cases for each error score. Error scores are based on converting status to a numeric scale (1 = *Red*, 5 = *Amber*) and calculating residuals (i.e. predicted - actual). A positive error score means that the algorithm assigned a better status than the expert process. These numbers match the previous table but provide more detail. For example, the sum of counts for all negative error scores in this table equals the number for *PredWorse* in the previous table.

ErrorScore	All	Chinook	Coho	Sockeye
-4	1	NA	NA	1
-2	6	2	NA	4
0	49	11	5	33
2	8	2	NA	6
Total	64	15	5	44

F.2 ERROR DIAGNOSTICS – FANCY PANTS

Table 31: Fancy Pants – Confusion Matrix. Layout as per Table 28.

Integrated Status

Predicted	*Red*	*Red/Amber*	*Amber*	*Amber/Green*	*Green*
Red	23	0	1	1	1
Red/Amber	0	5	0	0	1
Amber	0	0	8	1	0
Amber/Green	0	0	1	5	0
Green	0	0	0	1	6
None	2	2	3	2	2

Table 32: Fancy Pants – Error Type Frequency. Layout as per Table 29.

ErrorType	All	Chinook	Coho	Sockeye
NA	1	NA	NA	1
None	49	11	5	33
PredBetter	8	2	NA	6
PredWorse	7	2	NA	5
Total(excl.NA)	64	15	5	44
PercCorrect	77	73	100	75
PercWorse	11	13	NA	11
PercBetter	12	13	NA	14

Table 33: Fancy Pants – Error Scores Frequency. Layout as per Table 30.

ErrorScore	All	Chinook	Coho	Sockeye
-4	1	NA	NA	1
-3	1	NA	NA	1
-2	2	NA	NA	2
0	38	11	3	24
1	11	NA	2	9
2	1	NA	NA	1
Total	54	11	5	38

F.3 ERROR DIAGNOSTICS – CATEGORICAL REALIST

Table 34: Categorical Realist – Confusion Matrix. Layout as per Table 28.

Integrated Status

Predicted	*Red*	*Amber*	*Green*
Red	21	0	0
Amber	5	20	9
None	6	3	1

Table 35: Categorical Realist – Error Type Frequency. Layout as per Table 29.

ErrorType	**All**	**Chinook**	**Coho**	**Sockeye**
NA	10	NA	NA	10
None	41	10	5	26
PredBetter	5	3	NA	2
PredWorse	9	2	NA	7
Total(excl.NA)	55	15	5	35
PercCorrect	75	67	100	74
PercWorse	16	13	NA	20
PercBetter	9	20	NA	6

Table 36: Categorical Realist – Error Scores Frequency. Layout as per Table 30.

ErrorScore	**All**	**Chinook**	**Coho**	**Sockeye**
-2	9	2	NA	7
0	41	10	5	26
2	5	3	NA	2
Total	55	15	5	35

F.4 ERROR DIAGNOSTICS – SIMPLY *RED*

Table 37: Simply Red – Confusion Matrix. Layout as per Table 28.

	Integrated Status		
Predicted	*Red*	*Amber*	*Green*
Red	26	4	2
NotRed	2	15	6
None	4	4	2

Table 38: Simply Red – Error Type Frequency. Layout as per Table 29.

ErrorType	**All**	**Chinook**	**Coho**	**Sockeye**
NA	10	4	NA	6
None	26	10	NA	16
PredBetter	17	NA	5	12
PredWorse	12	1	NA	11
Total(excl.NA)	55	11	5	39
PercCorrect	47	91	NA	41
PercWorse	22	9	NA	28
PercBetter	31	NA	100	31

Table 39: Simply Red – Error Scores Frequency. Layout as per Table 30.

ErrorScore	**All**	**Chinook**	**Coho**	**Sockeye**
-4	2	NA	NA	2
-2	4	NA	NA	4
-1	6	1	NA	5
0	26	10	NA	16
1	15	NA	5	10
3	2	NA	NA	2
Total	55	11	5	39

F.5 ERROR DIAGNOSTICS – LEARNING TREE 1

Table 40: Learning Tree 1 – Confusion Matrix. Layout as per Table 28.

Integrated Status

Predicted	*Red*	*Amber*	*Green*
Red	26	0	1
Amber	4	12	1
Amber	2	11	8

Table 41: Learning Tree 1 – Error Type Frequency. Layout as per Table 29.

ErrorType	**All**	**Chinook**	**Coho**	**Sockeye**
NA	0	0	0	0
None	46	12	NA	34
PredBetter	17	3	5	9
PredWorse	2	NA	NA	2
Total(excl.NA)	65	15	5	45
PercCorrect	71	80	NA	76
PercWorse	3	NA	NA	4
PercBetter	26	20	100	20

Table 42: Learning Tree 1 – Error Scores Frequency. Layout as per Table 30.

ErrorScore	**All**	**Chinook**	**Coho**	**Sockeye**
-4	1	NA	NA	1
-2	1	NA	NA	1
0	46	12	NA	34
2	15	1	5	9
4	2	2	NA	NA
Total	65	15	5	45

F.6 ERROR DIAGNOSTICS – LEARNING TREE 2

Table 43: Learning Tree 2 – Confusion Matrix. Layout as per Table 28.

	Integrated Status		
Predicted	*Red*	*Amber*	*Green*
Red	25	0	0
Amber	6	14	1
Amber	1	9	9

Table 44: Learning Tree 2 – Error Type Frequency. Layout as per Table 29.

ErrorType	All	Chinook	Coho	Sockeye
NA	0	0	0	0
None	48	12	NA	36
PredBetter	16	3	5	8
PredWorse	1	NA	NA	1
Total(excl.NA)	65	15	5	45
PercCorrect	74	80	NA	80
PercWorse	2	NA	NA	2
PercBetter	25	20	100	18

Table 45: Learning Tree 2 – Error Scores Frequency. Layout as per Table 30.

ErrorScore	All	Chinook	Coho	Sockeye
-2	1	NA	NA	1
0	48	12	NA	36
2	15	2	5	8
4	1	1	NA	NA
Total	65	15	5	45

F.7 ERROR DIAGNOSTICS – LEARNING TREE 3

Table 46: Learning Tree 3 – Confusion Matrix. Layout as per Table 28.

Integrated Status

Predicted	*Red*	*Amber*	*Green*
Red	28	1	1
Amber	4	19	2
Amber	0	3	7

Table 47: Learning Tree 3 – Error Type Frequency. Layout as per Table 29.

ErrorType	**All**	**Chinook**	**Coho**	**Sockeye**
NA	0	0	0	0
None	54	12	4	38
PredBetter	7	2	1	4
PredWorse	4	1	NA	3
Total(excl.NA)	65	15	5	45
PercCorrect	83	80	80	84
PercWorse	6	7	NA	7
PercBetter	11	13	20	9

Table 48: Learning Tree 3 – Error Scores Frequency. Layout as per Table 30.

ErrorScore	**All**	**Chinook**	**Coho**	**Sockeye**
-4	1	NA	NA	1
-2	3	1	NA	2
0	54	12	4	38
2	7	2	1	4
Total	65	15	5	45

APPENDIX G: RETROSPECTIVE TEST – SUMMARY OF RESULTS

G.1 COMPLETION RATES AND AGREEMENT BETWEEN ALGORITHMS

The number of algorithms that could assign a status to an individual case depends on how many of the six standard metrics are available for that CU in that year (Table 49). All seven fitted and constructed algorithms could assign a rapid status to the 492 cases that had all six status metrics. However, there are pronounced differences in completion rate between algorithms (Table G.2).

Learning Trees 1,2 and *3* could assign a rapid status to all 639 cases that had four or more of the six standard WSP metrics, and 99%-100% of the cases with 2 metrics. *Learning Tree 3* could assign status to all 822 cases with two or more of the six standard metrics, for an overall completion rate of 74%. All other algorithms had low or 0% completion rate for some cases with two to five metrics (e.g. *Minimalist*, *Fancy Pants*, and *Simply Red* could not assign a status for any of the 32 cases with five metrics). The *Minimalist* and *Learning Tree 3* algorithms are the only ones that could assign status to all 183 cases with two metrics, but *Categorical Realist* and the first two *Learning Tree* variations could also assign a status for most of these data-limited cases. *Learning Tree 3* was able to classify all of the cases with 2 or more metrics.

We are still exploring alternative approaches for summarizing retrospective patterns in algorithm performance. One approach is to compare everything to one preferred or benchmark algorithm.

We used *Learning Tree 3* as the benchmark, because it outperformed the other candidate algorithms in terms of percent completed and percent correct in the Learning Data Set.

Table G.3 is a nested version of the standard confusion matrices in Appendix F, comparing the five other algorithms to *Learning Tree 3*. For cases where *Learning Tree 3* assigned *Red* status, most of the other algorithms also assigned *Red*, with the exception of the *CategoricalRealis*t. For cases where *Learning Tree 3* assigned *Amber*, four other algorithms also mostly assigned *Amber* (*CatReal,Minimalist, LearningTree1, LearningTree2*), while *FancyPants* statuses were evenly split across status categories from *Green* to *Red*.

Table 49: Summary of Retrospective Test-Completion Rate by number of metrics and number of algorithms. There are 860 individual cases where a WSP rapid status could be assigned across CUs and years by one or more of the seven candidate algorithms. For these CUs and years, values could be estimated for one or more metrics. As a reminder, the metrics include: long-term trend, percent change, relative abundance, and absolute abundance. Perhaps obvious, but if no metric values could be estimated for a year and CU, no algorithm could assign a WSP status; this occurred for 316 cases. Conversely, if all four metric values could be estimated for a year and CU, then all seven algorithms could assign statuses; this occurred for 509 cases. For number of metrics from 1 to 3, varying numbers of algorithms could assign status.

Number of Metrics	Number of Algorithms 0	1	3	4	5	6	7
0	316	0	0	0	0	0	0
1	9	8	8	0	0	1	4
2	2	0	3	91	73	2	55
3	0	0	0	75	6	25	0
4	0	0	0	0	0	0	509
Total	**327**	**8**	**11**	**166**	**79**	**28**	**568**

Table 50: Summary of Retrospective Test - Completion Rate by Algorithm. The first column shows the number of metrics and cases (Year by CU). Metrics include long-term trend, percent change, relative abundance, absolute abundance. It shows how these cases are distributed across each of the seven algorithms as a percentage. For example, when number of metrics is two, there are 226 cases. *Minimalist* could assign a rapid status to 58% of them, while *Fancy Pants* could assign rapid status to only 25% of them, and *Learning Tree 3* could assign rapid status to 99% of them.

Number		**Percentage**						
Metrics	**Cases**	**Mini-malist**	**Fancy Pants**	**Cat Real**	**Simply Red**	**LTree1**	**LTree2**	**LTree3**
0	316	0	0	0	0	0	0	0
1	30	17	17	43	17	43	40	43
2	226	58	25	97	26	98	98	99
3	106	69	29	31	24	100	100	100
4	509	100	100	100	100	100	100	100
Total number	**1187**	**61**	**51**	**65**	**50**	**72**	**72**	**72**

Table 51: Summary of Retrospective Test - Comparing Learning Tree 3 to other Algorithms. Table shows possible combinations of rapid statuses for the *Learning Tree 3* algorithm (*LTree3*) compared to the status assignment from other algorithms (*Other*). Remaining columns list the number of cases in each pair of statuses. For example, numbers in the row where *LTree3* and *Other* both have *A* indicate how many cases were classified as *Amber* by both algorithms, if *Learning Tree 3* assigned *Amber*. Rows showing cases where both algorithms agreed are shaded in gray.

LTree 3	Other	Mini-malist	Fancy Pants	Cat Real	Simply Red	LTree1	LTree2
NA	NA	335	335	327	335	335	335
	G						
	AG						
	A						
	RA						
	R			8			
G	NA	5	28		28		
	G	74	81			152	146
	AG		6		115		
	A	64	24	129			6
	RA		4				
	R	9	9	23	9		
AG	NA						
	G						
	AG						
	A						
	RA						
	R						
A	NA	83	167	40	173		
	G	59	70			189	180
	AG		42		217		
	A	254	96	351		252	261
	RA		25				
	R	45	41	50	51		
RA	NA						
	G						
	AG						
	A						
	RA						
	R						
R	NA	45	55	46	53	2	3
	G	4					2
	AG		1		1		
	A	49	7	151		9	10
	RA		10				
	R	161	186	62	205	248	244

G.2 CHANGES SINCE LAST INTEGRATED STATUS ASSESSMENT

Four integrated status assessments under the WSP have been completed. These assessments covered 47 CUs from three species of salmon. *Learning Tree 3*, the recommended algorithm, indicates changes in status for many of the CUs since their last formal integrated assessment, using available data up to 2018 or 2019, depending on the CU.

G.2.1 Interior Fraser Coho

Five CUs of Interior Fraser coho were assessed in 2015 (DFO 2015) using spawner data up to 2013.

Learning Tree 3 indicates status changes since then for two of them:

- Fraser Canyon coho status dropped from *Amber* to *Red* in 2015 but improved back to *Amber* in 2018.
- North Thompson coho status dropped from *Green* to *Amber* in 2015 but improved back to *Green* in 2018.

G.2.2 Fraser Sockeye

22 CUs of Fraser sockeye were assessed in 2011 (Grant and Pestal 2012), using spawner data up to 2010, and 23 CUs were assessed in 2017 (Grant et al. 2020), using spawner data up to 2015.

Learning Tree 3 indicates worsening status since then for 11 of them:

- *Green* to *Amber* (5): Chilko-S-ES, Francois-Fraser-S, Pitt-ES, Shuswap-ES, Shuswap-L
- *Amber* to *Red* (5): Chilliwack-ES, Kamloops-ES, Lillooet-Harrison-L, Nahatlatch-ES, North-Barriere-ES
- *Green* to *Red* (1): Harrison-River (changed to *Amber* in 2017, and to *Red* in 2019)

Learning Tree 3 indicates improved status for one of them:

- *Red* to *Green*: Harrison-DS-L (*Red* was assigned by *Learning Tree 3*; experts assigned integrated status of *Amber/Green* with data up to 2015; With one more year of data, *Learning Tree 3* started assigning *Green*).

Learning Tree 3 indicates changes in status for one of them:

- Nadina-Francois-ES changed from *Amber* to *Red* for one year in 2017, then changed back to *Amber*.

G.2.3 Southern BC Chinook

Integrated status assessments were completed for 15 CUs of Southern BC Chinook in 2012 (DFO 2016) using spawner data up to 2012.

Learning Tree 3 indicates worsening status since then for 4 of the 15 CUs:

- *Green* to *Amber* (1): Upper Fraser River Spring 1.3 (CK-12)
- *Amber* to *Red* (3): Lower Fraser River Fall 0.3 (CK-03), Middle Fraser River Portage Fall 1.3 (CK-09; *Learning Tree 3* assigned *Amber* for data up to 2012, but the expert workshop assigned *Red*, with one additional year of data the algorithm also assigned *Red*), Middle Fraser River Summer 1.3 (CK-11).

Learning Tree 3 indicates changes in status since then for 3 of them:

- Lower Thompson Spring 1.2 (CK-17) improved to *Amber* in 2014 but dropped back to *Red* in 2019.
- Middle Fraser River Spring 1.3 (CK-10) improved to *Amber* in 2014 but dropped back to *Red* in 2019.
- WCVI Nootka & Kyuquot Fall 0.x (CK-32) improved to *Green* in 2015 but dropped back to *Amber* in 2019.

APPENDIX H: RETROSPECTIVE TEST – DETAILED RESULTS BY CONSERVATION UNIT

H.1 OVERVIEW

Rapid statuses can differ between candidate algorithms. The rapid statuses can also change over time as the CU abundance changes and the WSP metrics change accordingly. The retrospective test showed that patterns differ between CUs, so this appendix includes the detailed retrospective results. For each CU, it shows the pattern of abundance and the corresponding pattern in rapid status, and any integrated statuses that have been completed. Results are grouped by species and area or timing group. We included a brief summary of observed patterns at the beginning of each section:

- Interior Fraser Coho (Appendix H.2)
- Fraser Sockeye - Early Stuart (Appendix H.3)
- Fraser Sockeye - Early Summer (Appendix H.4)
- Fraser Sockeye - Summer (Appendix H.5)
- Fraser Sockeye - Late (Appendix H.6)
- Fraser Sockeye - River-Type (Appendix H.7)
- Southern BC Chinook – Fraser - Lower (Appendix H.8)
- Southern BC Chinook – Fraser - Upper (Appendix H.9)
- Southern BC Chinook – Fraser - Thompson (Appendix H.10)
- Southern BC Chinook - Inner South Coast (Appendix H.11)
- Southern BC Chinook - West Coast Vancouver Island (Appendix H.12)

H.2 INTERIOR FRASER COHO

Full integrated status assessments of Interior Fraser coho were completed in 2015 (DFO 2015), using spawner data up to 2013. Available spawner estimates were reviewed back to 1998 for the integrated assessment, and the retrospective test of rapid status algorithms also excluded earlier data. The integrated assessment and retrospective test cover 5 CUs:

- *Fraser Canyon* (**Figure 32**): The integrated status assessment for 2013 was *Amber*. Five of the seven algorithms generate rapid statuses that are consistent with the integrated assessment for 2013 (4 *Amber*, 1 *NotRed*). Four of the seven algorithms could assign a rapid status back to 2000 and give stable statuses for 2000-2014. Six of the seven algorithms flag a deteriorating status for a few years starting in 2015, 2 years after the integrated assessment. For data up to 2013, *Learning Tree 3* assigns *Amber* and matches the integrated assessment.
- *Middle Fraser* (**Figure 33**): The integrated status assessment for 2013 was *Amber*. Five of the seven algorithms generate rapid statuses that are consistent with the integrated assessment for 2013 (4 *Amber*, 1 *NotRed*). Four of the seven algorithms could assign rapid statuses back to 1998, and most give stable statuses for long periods of time. Two of the algorithms flag a poorer status for 2006 to 2010. Six of the seven algorithms indicate a stable status since the integrated assessment was completed. For data up to 2013, *Learning Tree 3* assigns *Amber* and matches the integrated assessment.
- *Lower Thompson* (**Figure 34**): The integrated status assessment for 2013 was *Amber/Green*. Four of the seven algorithms generate rapid statuses that are consistent with the integrated assessment for 2013 (2 *Amber*, 1 *Amber/Green*, 1 *NotRed*). All seven algorithms could assign rapid statuses back to 2000, and most give stable statuses for long periods of time. Two of the seven algorithms flag a poorer status for earlier in the time series (up to 2007/2008). Five of the seven algorithms indicate a stable status since the integrated assessment was completed, but two algorithms indicate a deteriorating status (from *Green* or *Amber/Green* to *Amber*). For data up to 2013, *Learning Tree 3* assigns *Amber* and almost matches the integrated assessment of *Amber/Green*.
- *South Thompson* (**Figure 35**): The integrated status assessment for 2013 was *Amber*. Four of the seven algorithms generate rapid statuses that are consistent with the integrated assessment for 2013 (4 *Amber*, 1 *NotRed*). All seven algorithms could assign rapid statuses back to 2000. Only two of the seven algorithms indicate a stable status since the integrated assessment was completed. For data up to 2013, *Learning Tree 3* assigns *Amber* and matches the integrated assessment.
- *North Thompson* (**Figure 36**): The integrated status assessment for 2013 was *Amber/Green*. Four of the seven algorithms generate rapid statuses that are consistent with the integrated assessment for 2013 (2 *Amber*, 1 *Amber/Green*, 1 *NotRed*). All seven algorithms could assign rapid statuses back to 2000. Three of the seven algorithms indicate a stable status since the integrated assessment was completed. Four algorithms indicate a deteriorating status (from *Amber* or *Amber/Green* to *Red*) for 2015 to 2017. For data up to 2013, *Learning Tree 3* assigns *Green* and almost matches the integrated assessment of *Amber/Green*.

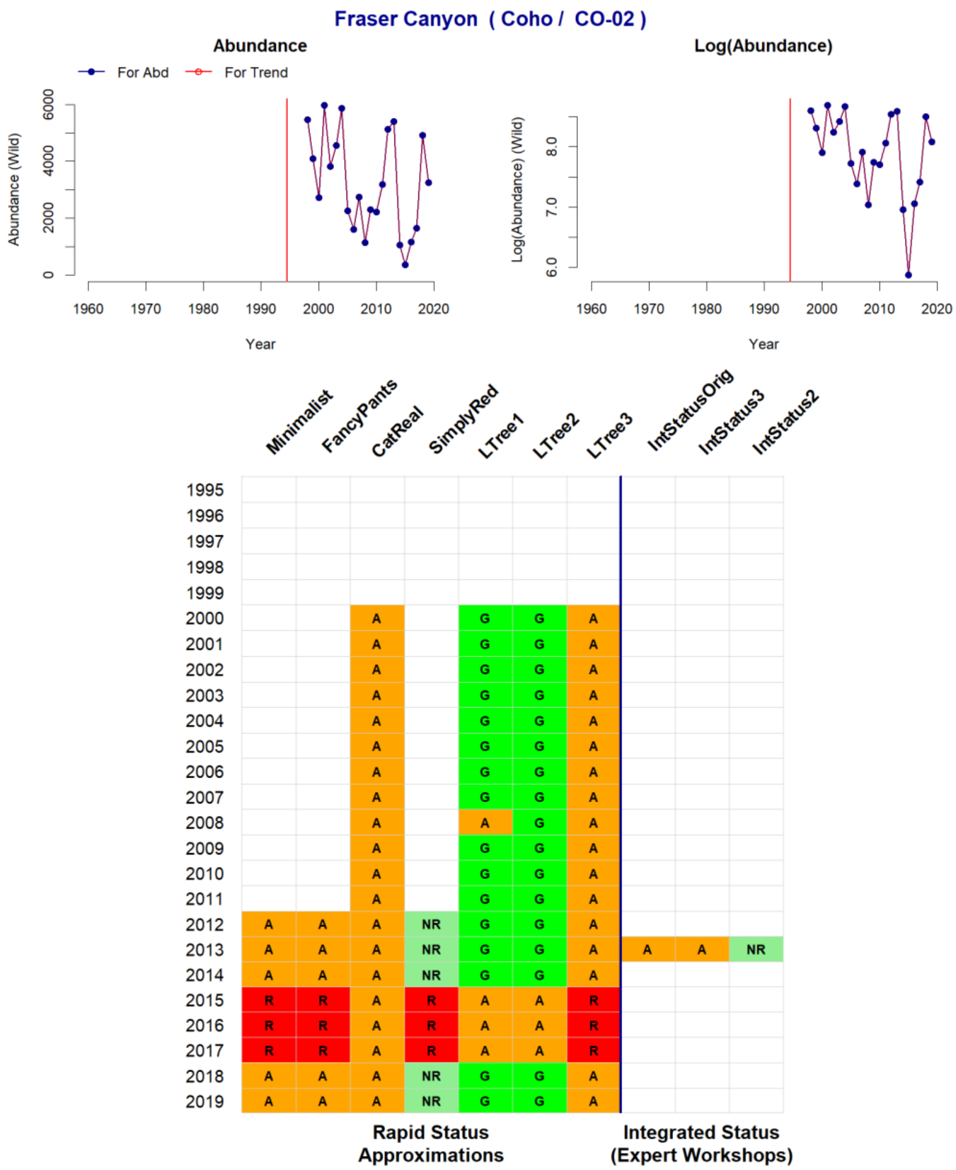


Figure 32. Retrospective test of rapid status - Coho - Fraser Canyon. Figure shows annual rapid statuses assigned by seven alternative candidate algorithms, as well as completed integrated status assessments. Note that integrated status assessments are mapped onto the year of data used, not the year the WSP status assessment was done. Statuses are denoted as R = *Red*, RA = *Red/Amber*, A = *Amber*, AG = *Amber/Green*, G = *Green*, NR = *NotRed*, DD = *Data Deficient*, and UD = *Undetermined*. For comparison, integrated statuses are shown as the original status assignment from the expert workshops (IntStatusOrig), as well as the conversions to the 3-status (*Red/Amber/Green*) and 2-status scales (*Red/NotRed*).

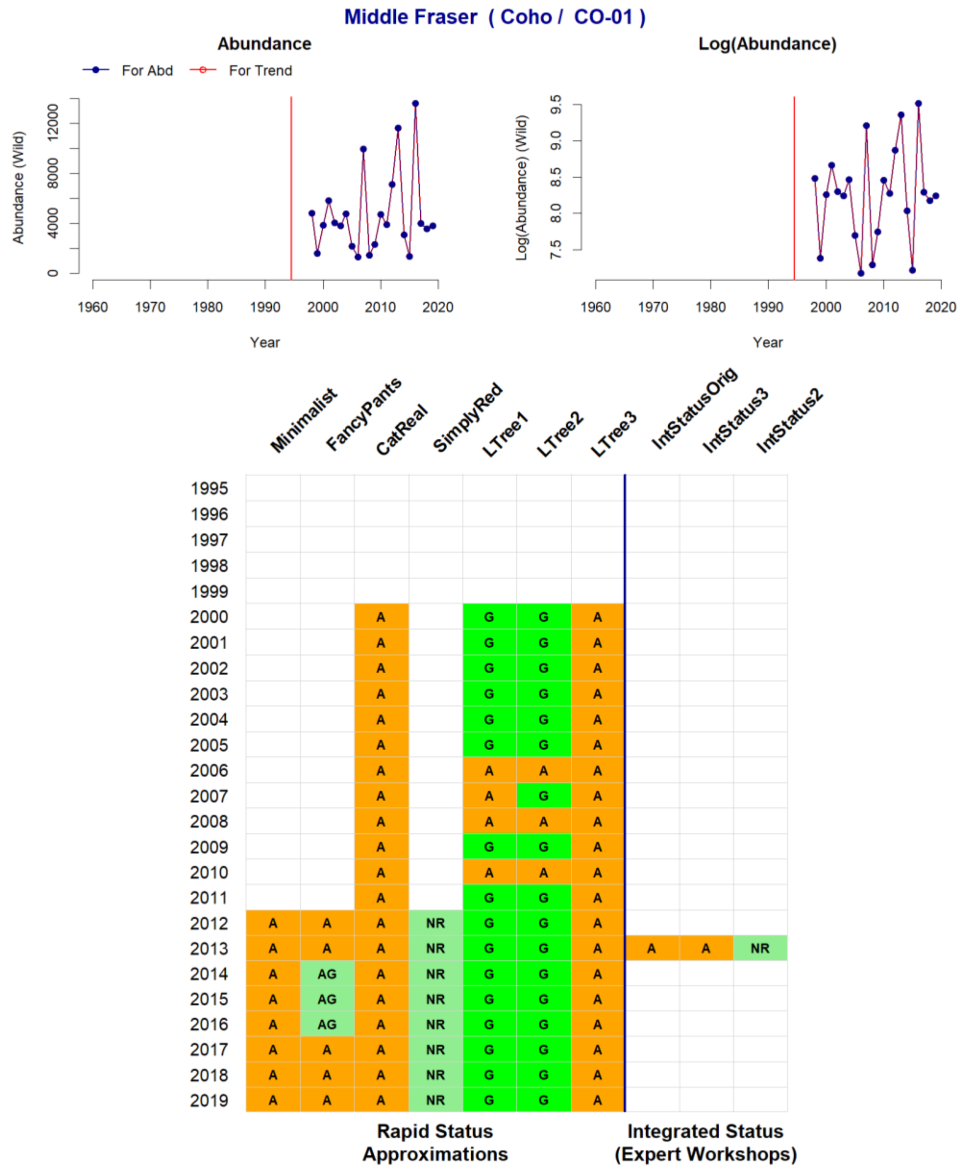


Figure 33. Retrospective test of rapid status - Coho – Middle Fraser. Figure shows annual rapid statuses assigned by seven alternative candidate algorithms, as well as completed integrated status assessments. Note that integrated status assessments are mapped onto the year of data used, not the year the WSP status assessment was done. Statuses are denoted as R = *Red*, RA = *Red/Amber*, A = *Amber*, AG = *Amber/Green*, G = *Green*, NR = *NotRed*, DD = *Data Deficient*, and UD = *Undetermined*. For comparison, integrated statuses are shown as the original status assignment from the expert workshops (IntStatusOrig), as well as the conversions to the 3-status (*Red/Amber/Green*) and 2-status scales (*Red/NotRed*).

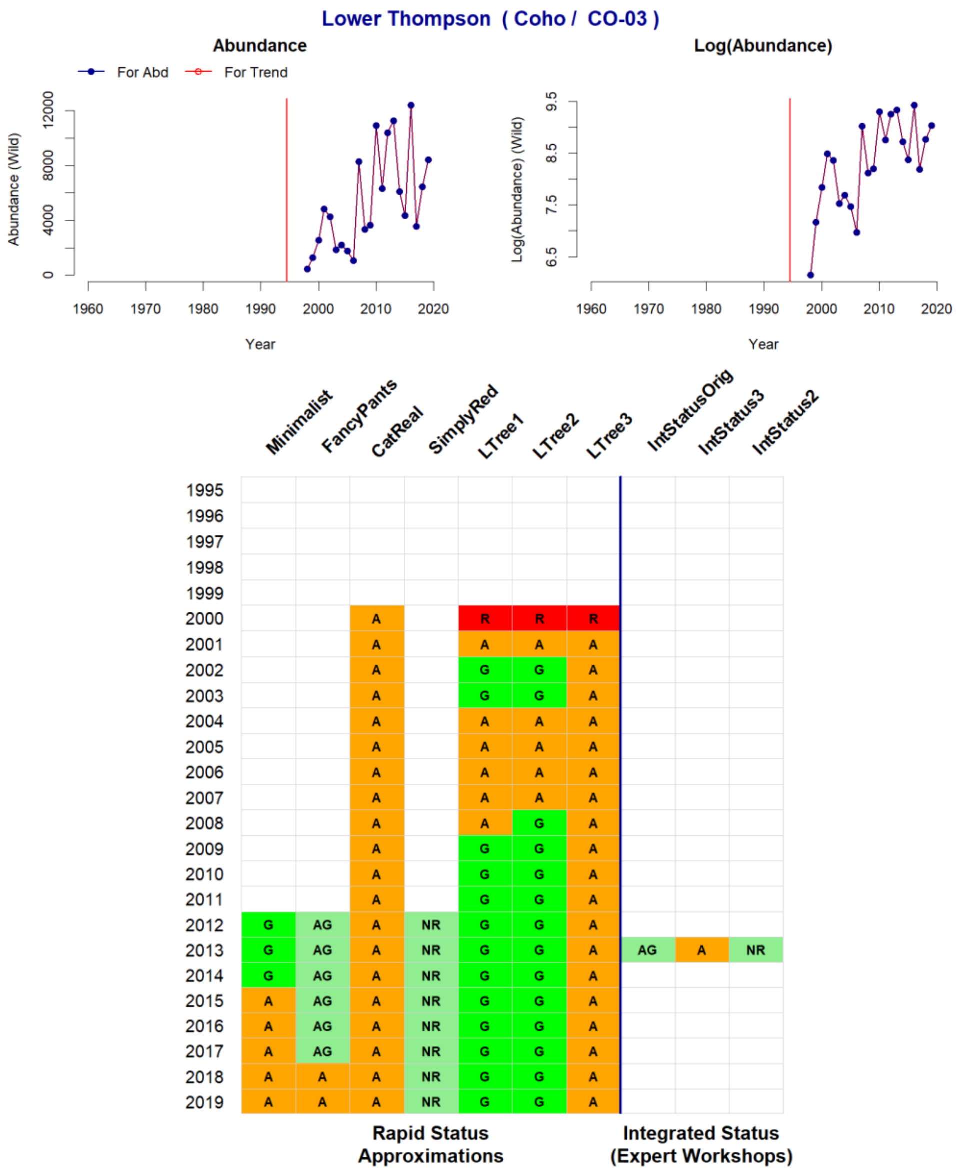
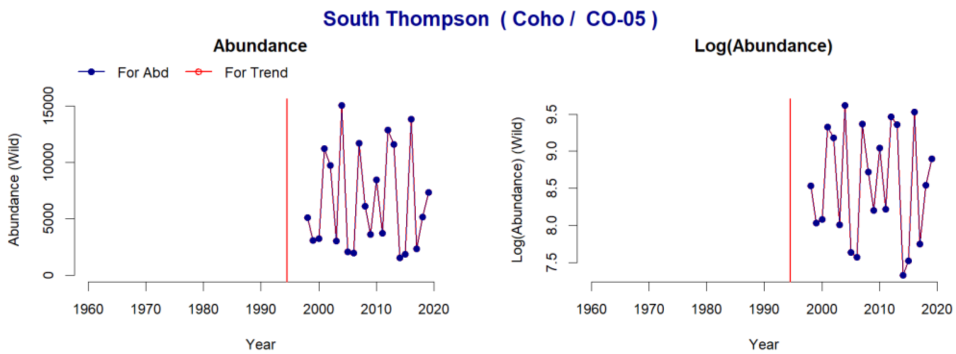


Figure 34. Retrospective test of rapid status - Coho – Lower Thompson. Figure shows annual rapid statuses assigned by seven alternative candidate algorithms, as well as completed integrated status assessments. Note that integrated status assessments are mapped onto the year of data used, not the year the WSP status assessment was done. Statuses are denoted as R = *Red*, RA = *Red/Amber*, A = *Amber*, AG = *Amber/Green*, G = *Green*, NR = *NotRed*, DD = *Data Deficient*, and UD = *Undetermined*. For comparison, integrated statuses are shown as the original status assignment from the expert workshops (IntStatusOrig), as well as the conversions to the 3-status (*Red/Amber/Green*) and 2-status scales (*Red/NotRed*).



	Minimalist	FancyPants	CatReal	SimplyRed	LTree1	LTree2	LTree3	IntStatusOrig	IntStatus3	IntStatus2
1995										
1996										
1997										
1998										
1999										
2000			A		A	A	A			
2001			A		A	G	A			
2002			A		G	G	A			
2003			A		G	G	A			
2004			A		G	G	A			
2005			A		A	G	A			
2006			A		A	A	A			
2007			A		A	A	A			
2008			A		G	G	A			
2009			A		G	G	A			
2010			A		G	G	A			
2011			A		A	G	A			
2012	A	A	A	NR	G	G	A			
2013	A	A	A	NR	G	G	A	A	A	NR
2014	A	AG	A	NR	G	G	A			
2015	R	R	A	R	A	A	A			
2016	R	R	A	R	A	A	A			
2017	R	R	A	R	A	A	A			
2018	A	A	A	NR	G	G	A			
2019	A	A	A	NR	A	A	A			

Rapid Status Approximations | Integrated Status (Expert Workshops)

\

Figure 35. Retrospective test of rapid status - Coho – South Thompson. Figure shows annual rapid statuses assigned by seven alternative candidate algorithms, as well as completed integrated status assessments. Note that integrated status assessments are mapped onto the year of data used, not the year the WSP status assessment was done. Statuses are denoted as R = *Red*, RA = *Red/Amber*, A = *Amber*, AG = *Amber/Green*, G = *Green*, NR = *NotRed*, DD = *Data Deficient*, and UD = *Undetermined*. For comparison, integrated statuses are shown as the original status assignment from the expert workshops (IntStatusOrig), as well as the conversions to the 3-status (*Red/Amber/Green*) and 2-status scales (*Red/NotRed*).

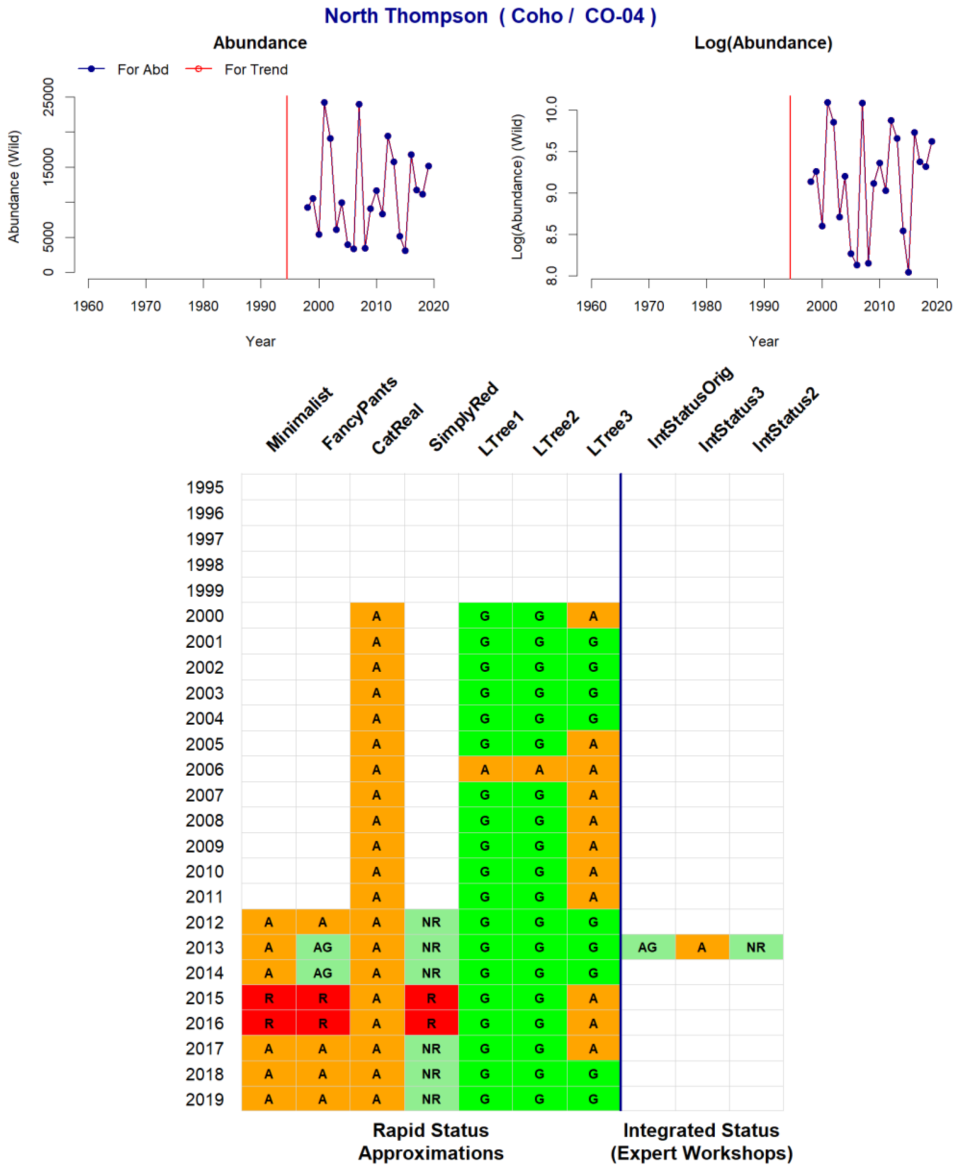


Figure 36. Retrospective test of rapid status - Coho – North Thompson. Figure shows annual rapid statuses assigned by seven alternative candidate algorithms, as well as completed integrated status assessments. Note that integrated status assessments are mapped onto the year of data used, not the year the WSP status assessment was done. Statuses are denoted as R = *Red*, RA = *Red/Amber*, A = *Amber*, AG = *Amber/Green*, G = *Green*, NR = *NotRed*, DD = *Data Deficient*, and UD = *Undetermined*. For comparison, integrated statuses are shown as the original status assignment from the expert workshops (IntStatusOrig), as well as the conversions to the 3-status (*Red/Amber/Green*) and 2-status scales (*Red/NotRed*).

H.3 FRASER SOCKEYE – EARLY STUART

Full integrated status assessments of Early Stuart sockeye were completed in 2011 (Grant and Pestal 2012) , using spawner data up to 2010, and in 2017 (Grant et al. 2020), using spawner data up to 2015. The integrated assessment and retrospective test covers 1 CU in the Early Stuart Sockeye management unit:

- *Takla-Trembleur-EStu* (**Figure 37**): The integrated status assessment was *Red* with data up to 2010 and *Red* with data up to 2015. All seven algorithms could assign statuses for every year since 1995. Six of the seven algorithms assigned *Green* or *NotRed* status at the beginning of the retrospective test (1995,1996), then indicated worsening status throughout the late 1990s and early 2000s, and then turned to *Red* around 2005. One algorithm (*Categorical realist*) assigned *Amber* status for the whole time series of the retrospective test. All algorithms indicate that status has not changed since the last integrated assessment using data up to 2015. For data up to 2010 and up to 2015, *Learning Tree 3* assigns *Red* and matches the integrated assessments.

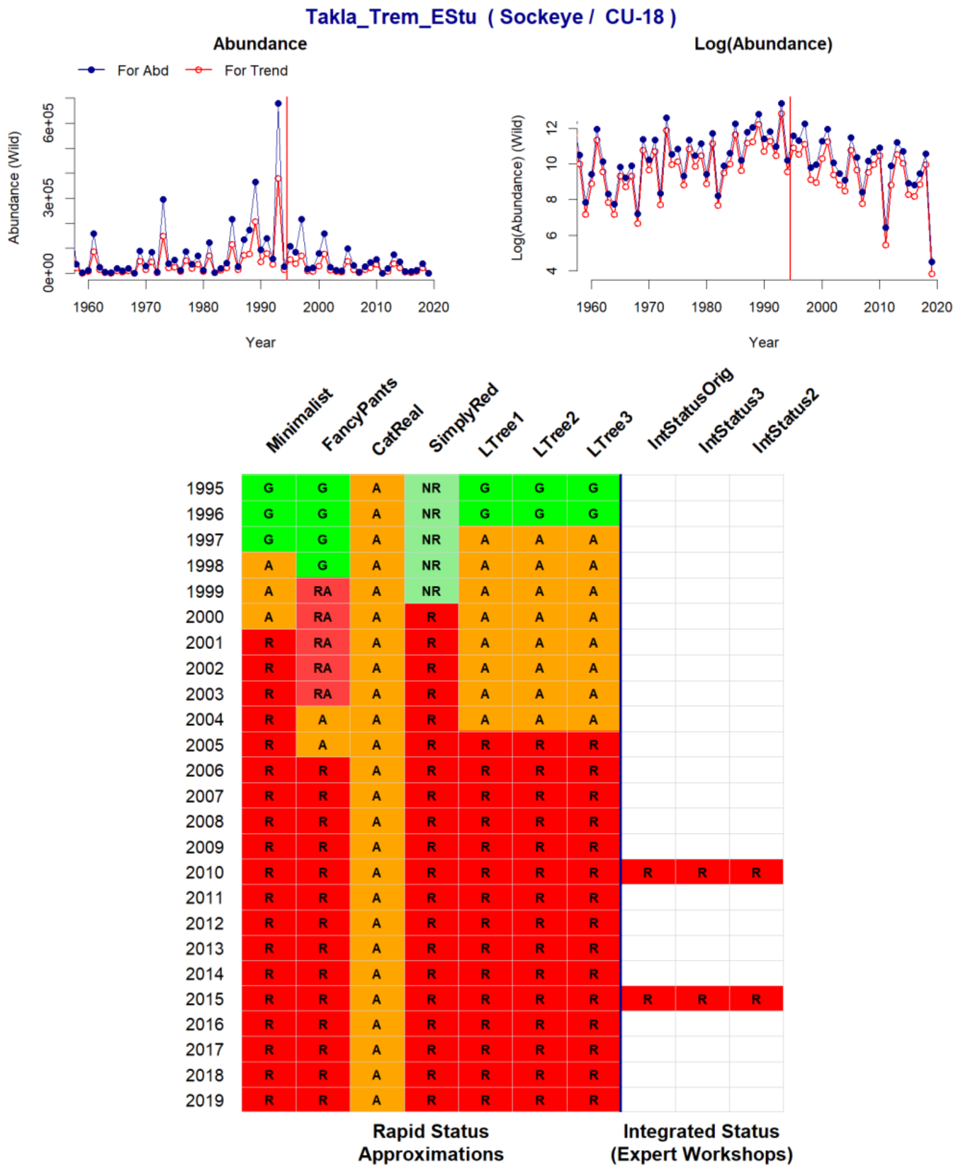


Figure 37. Retrospective test of rapid status - Fraser Sockeye - Takla_Trembleur_Early Stuart. Figure shows annual rapid statuses assigned by seven alternative candidate algorithms, as well as completed integrated status assessments. Note that integrated status assessments are mapped onto the year of data used, not the year the WSP status assessment was done. Statuses are denoted as R = *Red*, RA = *Red/Amber*, A = *Amber*, AG = *Amber/Green*, G = *Green*, NR = *NotRed*, DD = *Data Deficient*, and UD = *Undetermined*. For comparison, integrated statuses are shown as the original status assignment from the expert workshops (IntStatusOrig), as well as the conversions to the 3-status (*Red/Amber/Green*) and 2-status scales (*Red/NotRed*).

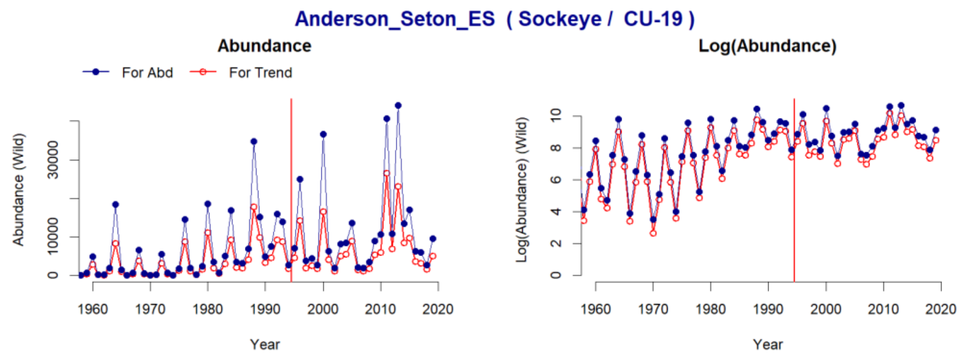
H.4 FRASER SOCKEYE – EARLY SUMMER

Full integrated status assessments of Early Summer sockeye were completed in 2011 (Grant and Pestal 2012) , using spawner data up to 2010, and in 2017 (Grant et al. 2020), using spawner data up to 2015. The integrated assessment and retrospective test covers 10 CUs in the Early Summer Sockeye management unit:

- *Anderson-Seton-ES* (**Figure 38**): The integrated status assessment was *Amber* with data up to 2010 and *Amber/Green* with data up to 2015. All seven algorithms could assign statuses for every year since 1995. All seven algorithms assigned either *Amber*, *Green*, *Amber/Green* or *NotRed* status to all years except 2009. 4 algorithms picked up a decline in status in 2009, but shifted back up to the previous status the next year, or shortly after. All algorithms indicate that status has not changed since the last integrated assessment using data up to 2015. For data up to 2010, *Learning Tree 3* assigns *Amber* and matches the integrated assessment. For data up to 2015, *Learning Tree 3* assigns *Amber* and almost matches integrated status: *Amber/Green*.
- *Bowron-ES* (**Figure 39**): The integrated status assessment was *Red* with data up to 2010 and *Red* with data up to 2015. All seven algorithms could assign statuses for every year since 1995. Six of the seven algorithms assigned *Red* status to all years since the early 2000s, the seventh assigned *Amber*. All algorithms indicate that status has not changed since the last integrated assessment using data up to 2015. For data up to 2010 and up to 2015, *Learning Tree 3* assigns *Red* matching integrated status.
- *Chilliwack-ES* (**Figure 40**): The integrated status assessment was *Red/Amber* with data up to 2010 and *Amber/Green* with data up to 2015. Four of the seven algorithms could assign statuses for every year since 2004. All seven algorithms could assign a status starting 2018. The three versions of the *Learning Tree* algorithm give identical statuses for every year. All three indicate that status was poorer from 2005-2011, then improved to *Amber* from 2012-2017, and shifted back to *Red* since 2018. For data up to 2010 and up to 2015, *Learning Tree 3* assigns *Red* and matches the integrated assessments. For data up to 2010, *Learning Tree 3* assigns *Red* and almost matches the integrated assessment of *Red/Amber*. For data up to 2015, *Learning Tree 3* assigns *Amber* and almost matches the integrated assessment of *Amber/Green*.
- *Kamloops-ES* (**Figure 41**): The integrated status assessment was *Amber* with data up to 2010 and *Amber* with data up to 2015. All seven algorithms could assign statuses for every year since 1995. Five of the seven algorithms picked up an improvement in status around the early 2000s. All seven algorithms assigned either *Amber* or *NotRed* status to all years from 2010 to 2018. Five of the seven algorithms indicate that status has worsened to *Red* since the last integrated assessment of *Amber* using data up to 2015. For data up to 2010 and up to 2015, *Learning Tree 3* assigns *Amber* and matches the integrated assessments.
- *Nadina/Francois-ES* (**Figure 42**): The integrated status assessment was *Red* with data up to 2010 and *Amber/Green* with data up to 2015. All seven algorithms could assign statuses for every year since 1995. All seven algorithms indicated consistent status from 1995 to 2014, but five of the algorithms produced changing annual statuses since 2015, switching between *Red* and either *Amber* or *NotRed*. Those five algorithms indicate that status improved in 2015 (which matches the integrated assessment), then worsened in 2017, and improved again to the previous *Amber* or *NotRed* since 2018. For data up to 2010, *Learning Tree 3* assigns *Red* and matches the integrated assessment. For data up to 2015, *Learning Tree 3* assigns *Amber* and

almost matches the integrated assessment of *Amber/Green*.

- *Nahatlatch-ES* (**Figure 43**): The integrated status assessment was *Red* with data up to 2010 and *Amber* with data up to 2015. Four of the seven algorithms could assign statuses for every year since 1995. Two additional algorithms could assign status for some years where the resulting status was *Red*. Four other algorithms also assigned *Red* status for those years. One algorithm, *Categorical Realist* could not assign status for any year. Algorithms that assigned statuses for all years generated *Amber* or *Green* statuses for 1995 to 2007, *Red* from 2008 to 2012 (matching the integrated expert assessment). From 2013 to 2018, all completed statuses point to an improvement (*Amber* or *Green* from the algorithms, *Amber* from the expert workshop), but all 6 completed statuses turned *Red* in 2019, indicating that status has worsened since the last WSP workshop. For data up to 2010*, Learning Tree 3* assigns *Red* and matches the integrated assessment. For data up to 2015, *Learning Tree 3* assigns *Amber* and matches the integrated assessment.
- *North Barriere-ES* (**Figure 44**): The integrated status assessment was *Amber* with data up to 2010 and *Amber* with data up to 2015. All seven algorithms could assign statuses for every year since 1995. Most algorithms indicate worse status in recent years compared to the late 1990s. All seven algorithms assigned *Amber* or *NotRed* for 2008 to 2017, matching the expert assessments of *Amber* with data up to 2010 and up to 2015. Four of the algorithms picked up a worsening status since 2018, shifting to *Red*. For data up to 2010 and up to 2015, *Learning Tree 3* assigns *Amber* and matches the integrated assessments.
- *Pitt-ES* (**Figure 45**): The integrated status assessment was *Amber/Green* with data up to 2010 and *Green* with data up to 2015. All seven algorithms could assign statuses for every year since 1995. Most of the algorithms give statuses that match the integrated assessment with data up to 2015, and almost match the integrated assessments with data up to 2010. The three versions of the *Learning Tree* algorithm assign *Green* to most years since 1995, but two of them switched to *Amber* in 2010, and all three switched to *Amber* in 2019. Overall, four of the seven algorithms indicate a worsening status in 2019.
- *Shuswap-ES* (**Figure 46**): The integrated status assessment was *Amber/Green* with data up to 2010 and *Amber* with data up to to 2015. All seven algorithms could assign statuses for every year since 1995. The three fitted algorithms (*Minimalist*, *Fancy Pants*, and *CategorcialRealist*) assigned *Amber* status for most years since 1995. The three versions of the *Learning Tree* algorithm switched from *Amber* to *Green* in 2002. Overall, four of the algorithms indicate a worsening status in 2019. For data up to 2010, *Learning Tree 3* assigns *Green* and almost matches the integrated assessment. For data up to 2015, *Learning Tree 3* assigns *Green*, a full status category better than the integrated assessment of *Amber*. Experts in the status reassessment workshop, looking at data up to 2015, used additional information to downgrade the status: (1) low abundance, (2) declining trends on off-cycle years.
- *Taseko-ES* (**Figure 47**): The integrated status assessment was *Red* with data up to 2010 and *Red* with data up to 2015. All seven algorithms could assign statuses for every year since 2006, but most algorithms have a 3-4 year gap in status assignments in the early 2000s. Almost all completed status assignments since 1997 are *Red*, except for 2005, where two algorithms assigned *Green*, two assigned *Amber*, one assigned *Red*, and two could not assign a status. For data up to 2010 and up to 2015, *Learning Tree 3* assigns *Red* and matches the integrated assessments.



Year	Minimalist	FancyPants	CatReal	SimplyRed	LTree1	LTree2	LTree3	IntStatusOrig	IntStatus3	IntStatus2
1995	G	AG	A	NR	A	A	A			
1996	G	AG	A	NR	A	A	A			
1997	G	A	A	NR	A	A	A			
1998	G	A	A	NR	A	A	A			
1999	G	A	A	NR	A	A	A			
2000	G	A	A	NR	A	A	A			
2001	G	A	A	NR	A	A	A			
2002	G	A	A	NR	A	A	A			
2003	G	A	A	NR	A	A	A			
2004	G	A	A	NR	A	A	A			
2005	G	A	A	NR	A	A	A			
2006	G	A	A	NR	A	A	A			
2007	G	A	A	NR	A	A	A			
2008	G	A	A	NR	A	A	A			
2009	A	A	A	R	R	R	R			
2010	A	A	A	NR	A	A	A	A	A	NR
2011	A	A	A	NR	A	A	A			
2012	G	A	A	NR	A	A	A			
2013	G	AG	A	NR	A	A	A			
2014	G	AG	A	NR	A	G	A			
2015	G	AG	A	NR	A	A	A	AG	A	NR
2016	G	AG	A	NR	A	A	A			
2017	G	AG	A	NR	A	A	A			
2018	G	AG	A	NR	A	A	A			
2019	G	AG	A	NR	A	A	A			

Rapid Status Approximations | Integrated Status (Expert Workshops)

Figure 38. Retrospective test of rapid status - Fraser Sockeye - Anderson-Seton-ES. Figure shows annual rapid statuses assigned by seven alternative candidate algorithms, as well as completed integrated status assessments. Note that integrated status assessments are mapped onto the year of data used, not the year the WSP status assessment was done. Statuses are denoted as R = *Red*, RA = *Red/Amber*, A = *Amber*, AG = *Amber/Green*, G = *Green*, NR = *NotRed*, DD = *Data Deficient*, and UD = *Undetermined*. For comparison, integrated statuses are shown as the original status assignment from the expert workshops (IntStatusOrig), as well as the conversions to the 3-status (*Red/Amber/Green*) and 2-status scales (*Red/NotRed*).

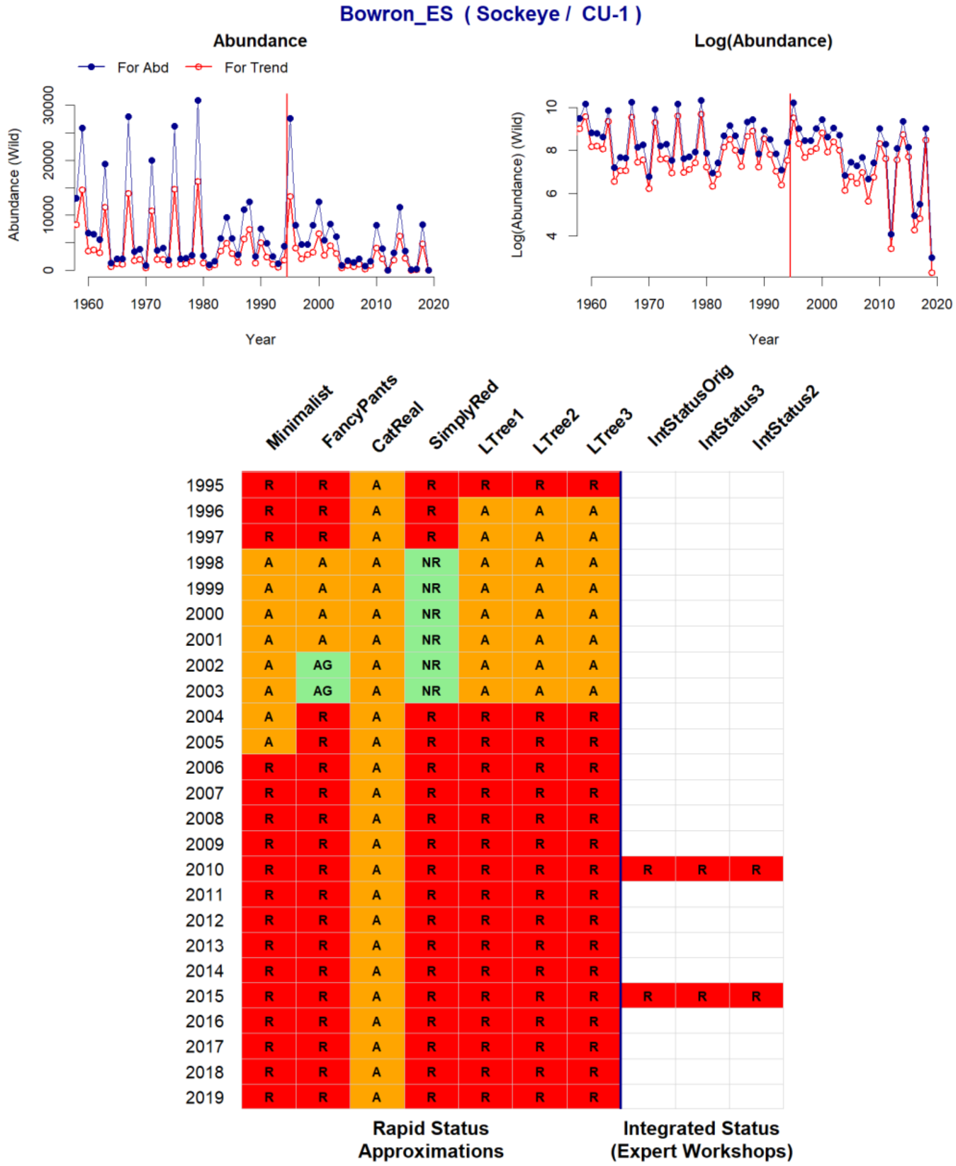


Figure 39. Retrospective test of rapid status - Fraser Sockeye - Bowron-ES. Figure shows annual rapid statuses assigned by seven alternative candidate algorithms, as well as completed integrated status assessments. Note that integrated status assessments are mapped onto the year of data used, not the year the WSP status assessment was done. Statuses are denoted as R = *Red*, RA = *Red/Amber*, A = *Amber*, AG = *Amber/Green*, G = *Green*, NR = *NotRed*, DD = *Data Deficient*, and UD = *Undetermined*. For comparison, integrated statuses are shown as the original status assignment from the expert workshops (IntStatusOrig), as well as the conversions to the 3-status (*Red/Amber/Green*) and 2-status scales (*Red/NotRed*).

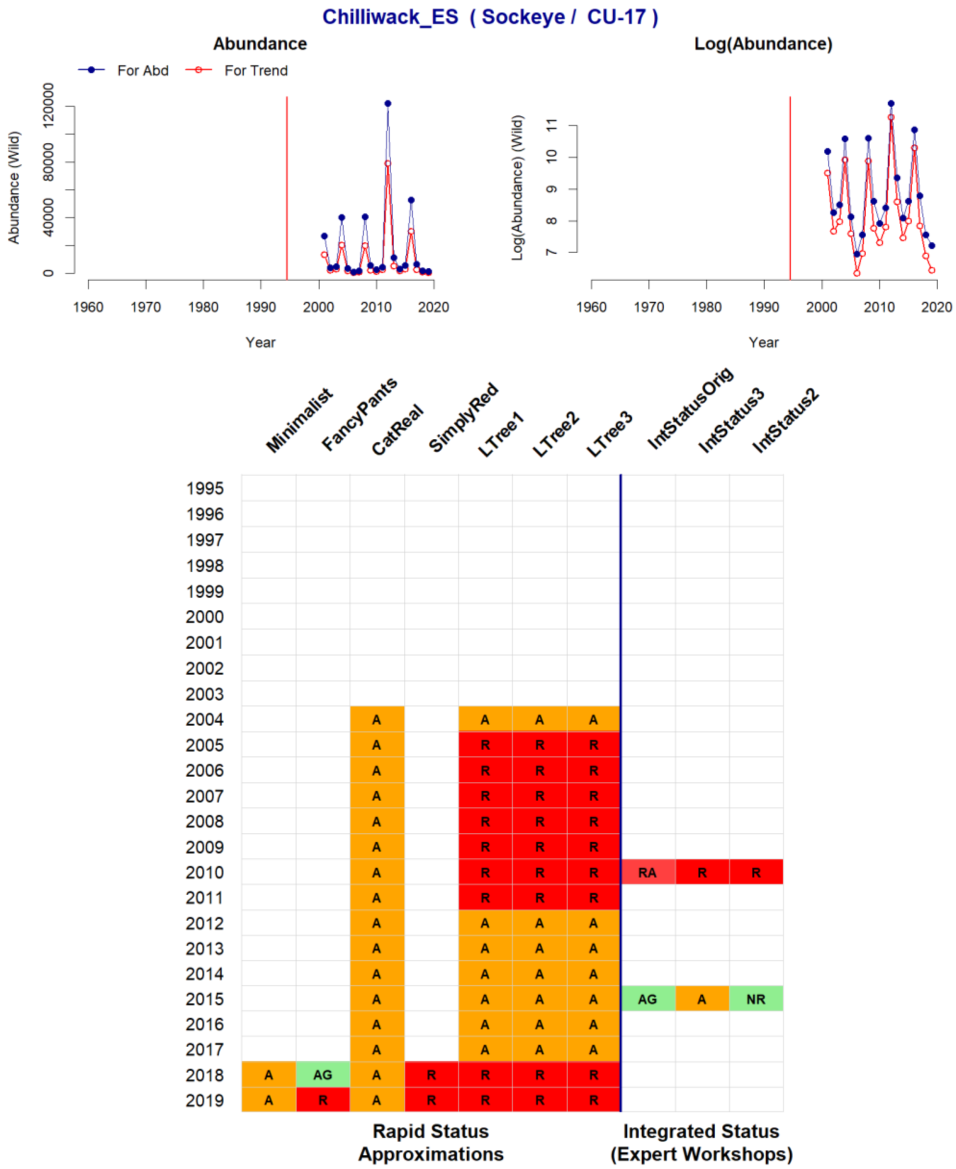
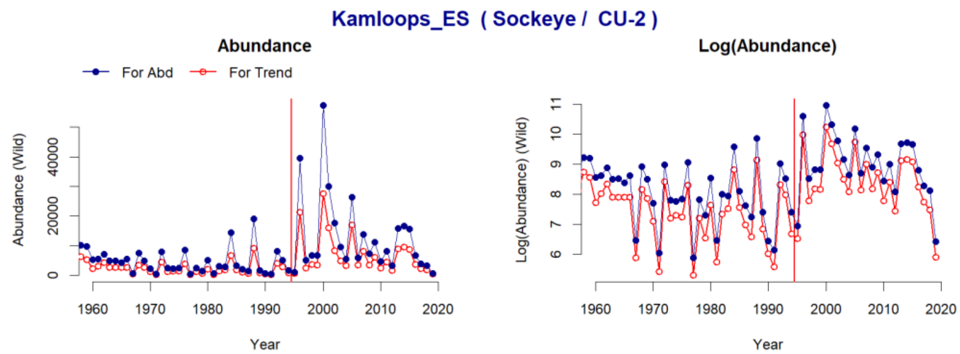


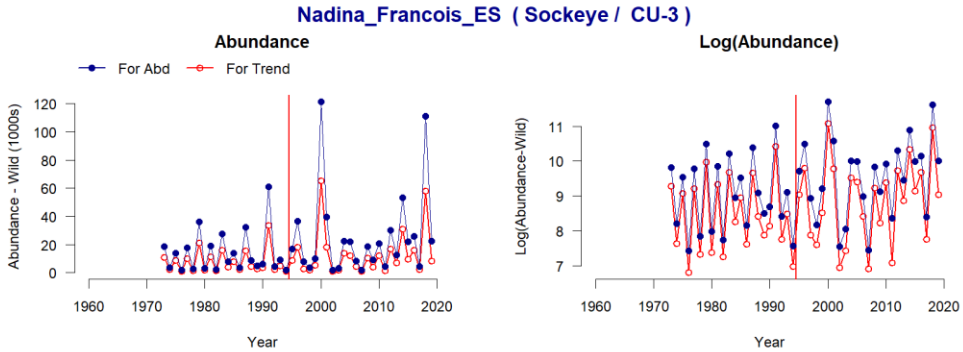
Figure 40. Retrospective test of rapid status - Fraser Sockeye - Chilliwack-ES. Figure shows annual rapid statuses assigned by seven alternative candidate algorithms, as well as completed integrated status assessments. Note that integrated status assessments are mapped onto the year of data used, not the year the WSP status assessment was done. Statuses are denoted as R = *Red*, RA = *Red/Amber*, A = *Amber*, AG = *Amber/Green*, G = *Green*, NR = *NotRed*, DD = *Data Deficient*, and UD = *Undetermined*. For comparison, integrated statuses are shown as the original status assignment from the expert workshops (IntStatusOrig), as well as the conversions to the 3-status (*Red/Amber/Green*) and 2-status scales (*Red/NotRed*).



	Minimalist	FancyPants	CatReal	SimplyRed	LTree1	LTree2	LTree3	IntStatusOrig	IntStatus3	IntStatus2
1995	R	R	A	R	R	R	R			
1996	A	R	A	R	R	R	R			
1997	A	R	A	R	R	R	R			
1998	A	A	A	NR	A	A	A			
1999	A	AG	A	NR	A	A	A			
2000	A	AG	A	NR	A	A	A			
2001	G	AG	A	NR	A	A	A			
2002	G	AG	A	NR	G	G	G			
2003	G	AG	A	NR	G	G	G			
2004	G	AG	A	NR	A	A	A			
2005	G	AG	A	NR	A	A	A			
2006	G	AG	A	NR	A	A	A			
2007	G	AG	A	NR	A	A	A			
2008	G	AG	A	NR	A	A	A			
2009	G	A	A	NR	A	A	A			
2010	A	A	A	NR	A	A	A	A	A	NR
2011	A	A	A	NR	A	A	A			
2012	A	A	A	NR	A	A	A			
2013	A	A	A	NR	A	A	A			
2014	A	A	A	NR	A	A	A			
2015	A	A	A	NR	A	A	A	A	A	NR
2016	A	A	A	NR	A	A	A			
2017	A	A	A	NR	A	A	A			
2018	A	A	A	NR	A	A	A			
2019	A	R	A	R	R	R	R			

Rapid Status Approximations — Integrated Status (Expert Workshops)

Figure 41. Retrospective test of rapid status - Fraser Sockeye - Kamloops-ES. Figure shows annual rapid statuses assigned by seven alternative candidate algorithms, as well as completed integrated status assessments. Note that integrated status assessments are mapped onto the year of data used, not the year the WSP status assessment was done. Statuses are denoted as R = *Red*, RA = *Red/Amber*, A = *Amber*, AG = *Amber/Green*, G = *Green*, NR = *NotRed*, DD = *Data Deficient*, and UD = *Undetermined*. For comparison, integrated statuses are shown as the original status assignment from the expert workshops (IntStatusOrig), as well as the conversions to the 3-status (*Red/Amber/Green*) and 2-status scales (*Red/NotRed*).



	Minimalist	FancyPants	CatReal	SimplyRed	LTree1	LTree2	LTree3	IntStatusOrig	IntStatus3	IntStatus2
1995	A	R	A	R	R	R	R			
1996	A	R	A	R	R	R	R			
1997	A	R	A	R	R	R	R			
1998	A	R	A	R	R	R	R			
1999	A	R	A	R	R	R	R			
2000	A	R	A	R	R	R	R			
2001	A	A	A	R	R	R	R			
2002	A	R	A	R	R	R	R			
2003	A	R	A	R	R	R	R			
2004	A	R	A	R	R	R	R			
2005	A	R	A	R	R	R	R			
2006	A	R	A	R	R	R	R			
2007	A	R	A	R	R	R	R			
2008	A	R	A	R	R	R	R			
2009	A	R	A	R	R	R	R			
2010	A	R	A	R	R	R	R	R	R	R
2011	A	R	A	R	R	R	R			
2012	A	R	A	R	R	R	R			
2013	A	R	A	R	R	R	R			
2014	A	R	A	R	R	R	R			
2015	A	AG	A	NR	A	A	A	AG	A	NR
2016	A	AG	A	NR	A	A	A			
2017	A	R	A	R	R	R	R			
2018	A	AG	A	NR	A	A	A			
2019	A	AG	A	NR	A	A	A			

Rapid Status Approximations

Integrated Status (Expert Workshops)

Figure 42. Retrospective test of rapid status - Fraser Sockeye - Nadina_Francois-ES. Figure shows annual rapid statuses assigned by seven alternative candidate algorithms, as well as completed integrated status assessments. Note that integrated status assessments are mapped onto the year of data used, not the year the WSP status assessment was done. Statuses are denoted as R = *Red*, RA = *Red/Amber*, A = *Amber*, AG = *Amber/Green*, G = *Green*, NR = *NotRed*, DD = *Data Deficient*, and UD = *Undetermined*. For comparison, integrated statuses are shown as the original status assignment from the expert workshops (IntStatusOrig), as well as the conversions to the 3-status (*Red/Amber/Green*) and 2-status scales (*Red/NotRed*).

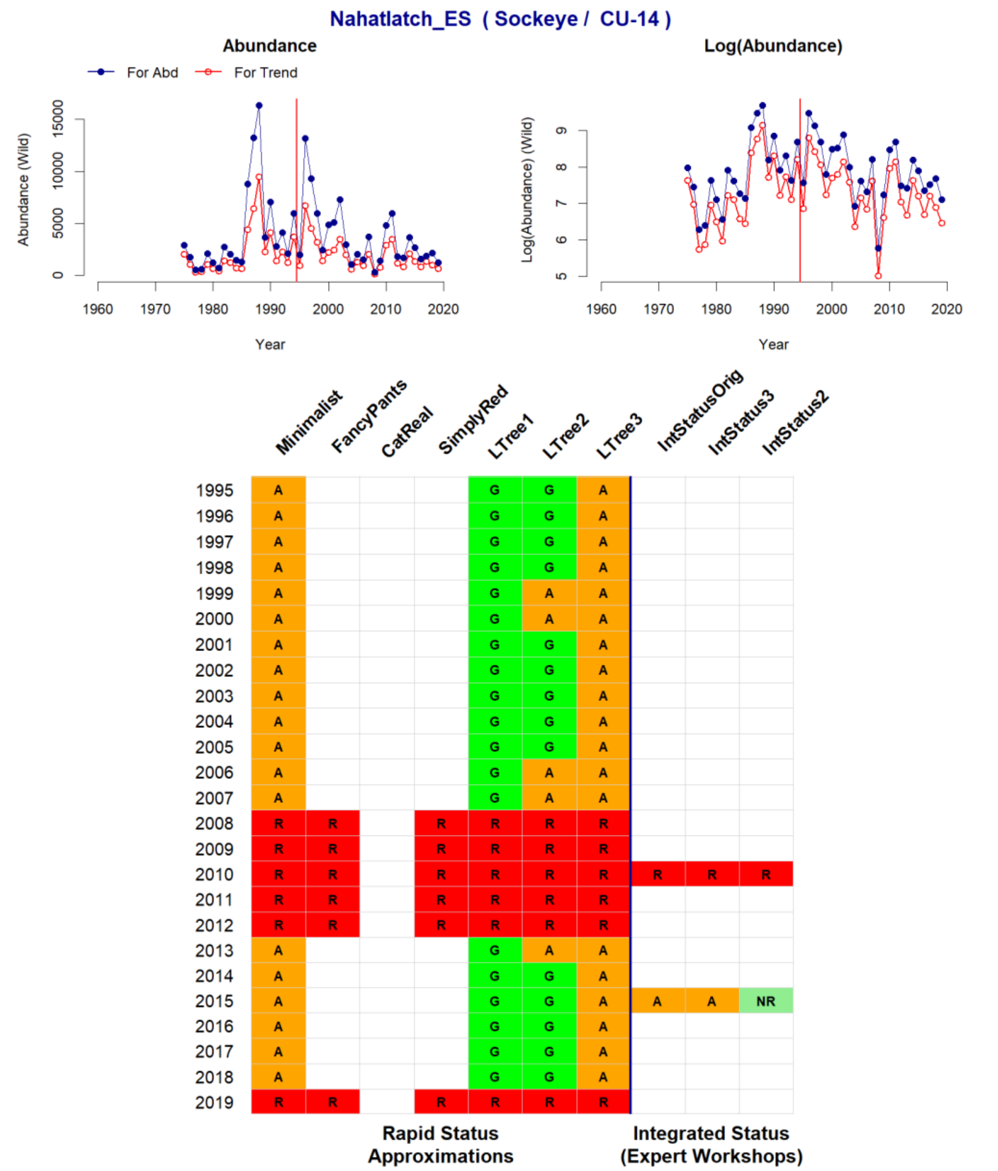
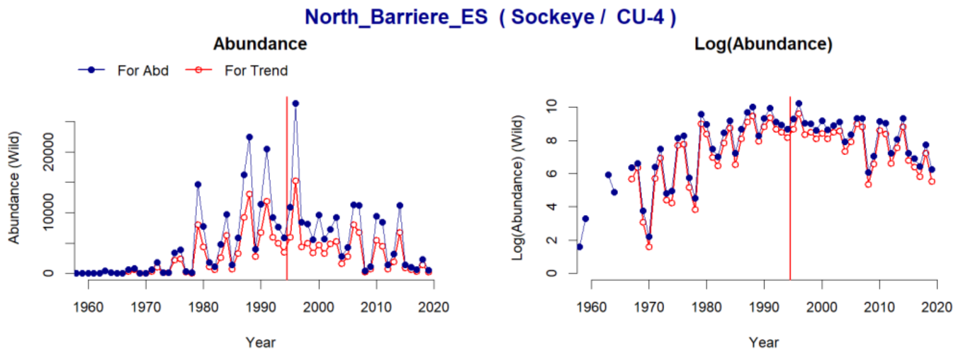


Figure 43. Retrospective test of rapid status – Fraser Sockeye - Nahatlatch-ES. Figure shows annual rapid statuses assigned by seven alternative candidate algorithms, as well as completed integrated status assessments. Note that integrated status assessments are mapped onto the year of data used, not the year the WSP status assessment was done. Statuses are denoted as R = *Red*, RA = *Red/Amber*, A = *Amber*, AG = *Amber/Green*, G = *Green*, NR = *NotRed*, DD = *Data Deficient*, and UD = *Undetermined*. For comparison, integrated statuses are shown as the original status assignment from the expert workshops (IntStatusOrig), as well as the conversions to the 3-status (*Red/Amber/Green*) and 2-status scales (*Red/NotRed*).

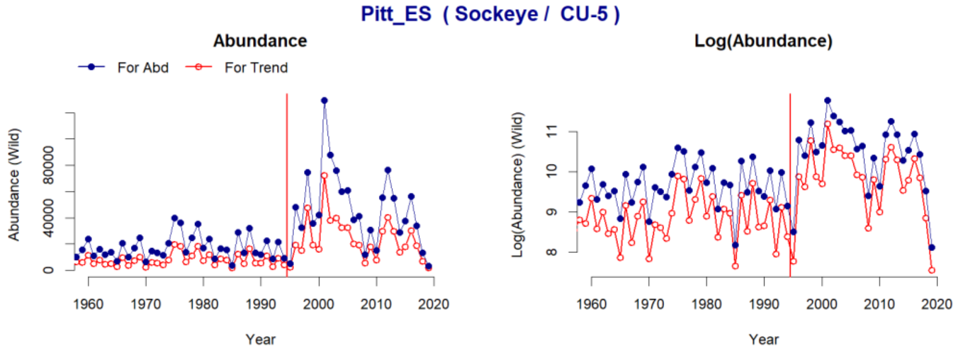


	Minimalist	FancyPants	CatReal	SimplyRed	LTree1	LTree2	LTree3	IntStatusOrig	IntStatus3	IntStatus2
1995	G	AG	A	NR	G	G	A			
1996	G	AG	A	NR	G	G	G			
1997	G	AG	A	NR	G	G	G			
1998	G	A	A	NR	G	G	G			
1999	G	A	A	NR	G	G	G			
2000	G	A	A	NR	G	G	A			
2001	G	A	A	NR	G	G	A			
2002	G	A	A	NR	G	G	A			
2003	G	A	A	NR	G	G	A			
2004	G	A	A	NR	G	G	A			
2005	A	A	A	NR	A	G	A			
2006	A	A	A	NR	G	G	A			
2007	A	A	A	NR	G	G	A			
2008	A	A	A	NR	A	A	A			
2009	A	A	A	NR	A	A	A			
2010	A	A	A	NR	A	A	A	A	A	NR
2011	A	A	A	NR	A	A	A			
2012	A	A	A	NR	A	A	A			
2013	A	A	A	NR	A	A	A			
2014	A	A	A	NR	A	A	A			
2015	A	A	A	NR	A	A	A	A	A	NR
2016	A	A	A	NR	A	A	A			
2017	A	A	A	NR	A	A	A			
2018	R	R	A	R	A	A	R			
2019	R	R	A	R	A	A	R			

Rapid Status Approximations

Integrated Status (Expert Workshops)

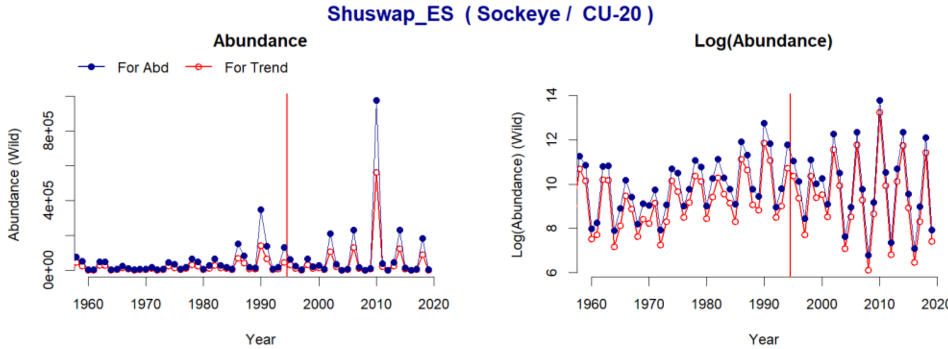
Figure 44. Retrospective test of rapid status - Fraser Sockeye - NorthBarriere-ES. Figure shows annual rapid statuses assigned by seven alternative candidate algorithms, as well as completed integrated status assessments. Note that integrated status assessments are mapped onto the year of data used, not the year the WSP status assessment was done. Statuses are denoted as R = *Red*, RA = *Red/Amber*, A = *Amber*, AG = *Amber/Green*, G = *Green*, NR = *NotRed*, DD = *Data Deficient*, and UD = *Undetermined*. For comparison, integrated statuses are shown as the original status assignment from the expert workshops (IntStatusOrig), as well as the conversions to the 3-status (*Red/Amber/Green*) and 2-status scales (*Red/NotRed*).



	Minimalist	FancyPants	CatReal	SimplyRed	LTree1	LTree2	LTree3	IntStatusOrig	IntStatus3	IntStatus2
1995	R	R	A	R	A	A	A			
1996	R	R	A	R	A	A	A			
1997	A	A	A	NR	A	A	A			
1998	A	A	A	NR	G	G	G			
1999	A	G	A	NR	G	G	G			
2000	A	G	A	NR	G	G	G			
2001	G	G	A	NR	G	G	G			
2002	G	G	A	NR	G	G	G			
2003	G	G	A	NR	G	G	G			
2004	G	G	A	NR	G	G	G			
2005	G	G	A	NR	G	G	G			
2006	G	G	A	NR	G	G	G			
2007	G	G	A	NR	G	G	G			
2008	G	G	A	NR	G	G	G			
2009	A	A	A	NR	G	G	G			
2010	A	A	A	NR	A	G	A	AG	A	NR
2011	A	A	A	NR	G	G	G			
2012	A	RA	A	NR	G	G	G			
2013	A	RA	A	NR	G	G	G			
2014	A	G	A	NR	G	G	G			
2015	A	G	A	NR	G	G	G	G	G	NR
2016	A	G	A	NR	G	G	G			
2017	A	G	A	NR	G	G	G			
2018	A	G	A	NR	G	G	G			
2019	A	A	A	NR	A	A	A			

Rapid Status Approximations — Integrated Status (Expert Workshops)

Figure 45. Retrospective test of rapid status - Fraser Sockeye – Pitt-ES. Figure shows annual rapid statuses assigned by seven alternative candidate algorithms, as well as completed integrated status assessments. Note that integrated status assessments are mapped onto the year of data used, not the year the WSP status assessment was done. Statuses are denoted as R = *Red*, RA = *Red/Amber*, A = *Amber*, AG = *Amber/Green*, G = *Green*, NR = *NotRed*, DD = *Data Deficient*, and UD = *Undetermined*. For comparison, integrated statuses are shown as the original status assignment from the expert workshops (IntStatusOrig), as well as the conversions to the 3-status (*Red/Amber/Green*) and 2-status scales (*Red/NotRed*).



	Minimalist	FancyPants	CatReal	SimplyRed	LTree1	LTree2	LTree3	IntStatusOrig	IntStatus3	IntStatus2
1995	A	G	A	NR	A	A	A			
1996	A	G	A	NR	A	A	A			
1997	A	G	A	NR	A	A	A			
1998	A	A	A	NR	A	A	A			
1999	A	A	A	NR	A	A	A			
2000	A	A	A	NR	A	A	A			
2001	A	A	A	NR	A	A	A			
2002	A	RA	A	NR	G	G	G			
2003	A	G	A	NR	G	G	G			
2004	A	A	A	NR	G	G	G			
2005	A	A	A	NR	G	G	G			
2006	A	A	A	NR	G	G	G			
2007	A	A	A	NR	G	G	G			
2008	A	A	A	NR	G	G	G			
2009	A	A	A	NR	G	G	G			
2010	A	A	A	NR	G	G	G	AG	A	NR
2011	A	A	A	NR	G	G	G			
2012	A	A	A	NR	G	G	G			
2013	A	G	A	NR	G	G	G			
2014	A	A	A	NR	G	G	G			
2015	A	A	A	NR	G	G	G	A	A	NR
2016	A	A	A	NR	G	G	G			
2017	A	A	A	NR	G	G	G			
2018	A	A	A	NR	G	G	G			
2019	R	R	A	R	G	G	A			

Rapid Status Approximations — Integrated Status (Expert Workshops)

Figure 46. Retrospective test of rapid status - Fraser Sockeye – Shuswap-ES. Figure shows annual rapid statuses assigned by seven alternative candidate algorithms, as well as completed integrated status assessments. Note that integrated status assessments are mapped onto the year of data used, not the year the WSP status assessment was done. Statuses are denoted as R = *Red*, RA = *Red/Amber*, A = *Amber*, AG = *Amber/Green*, G = *Green*, NR = *NotRed*, DD = *Data Deficient*, and UD = *Undetermined*. For comparison, integrated statuses are shown as the original status assignment from the expert workshops (IntStatusOrig), as well as the conversions to the 3-status (*Red/Amber/Green*) and 2-status scales (*Red/NotRed*).

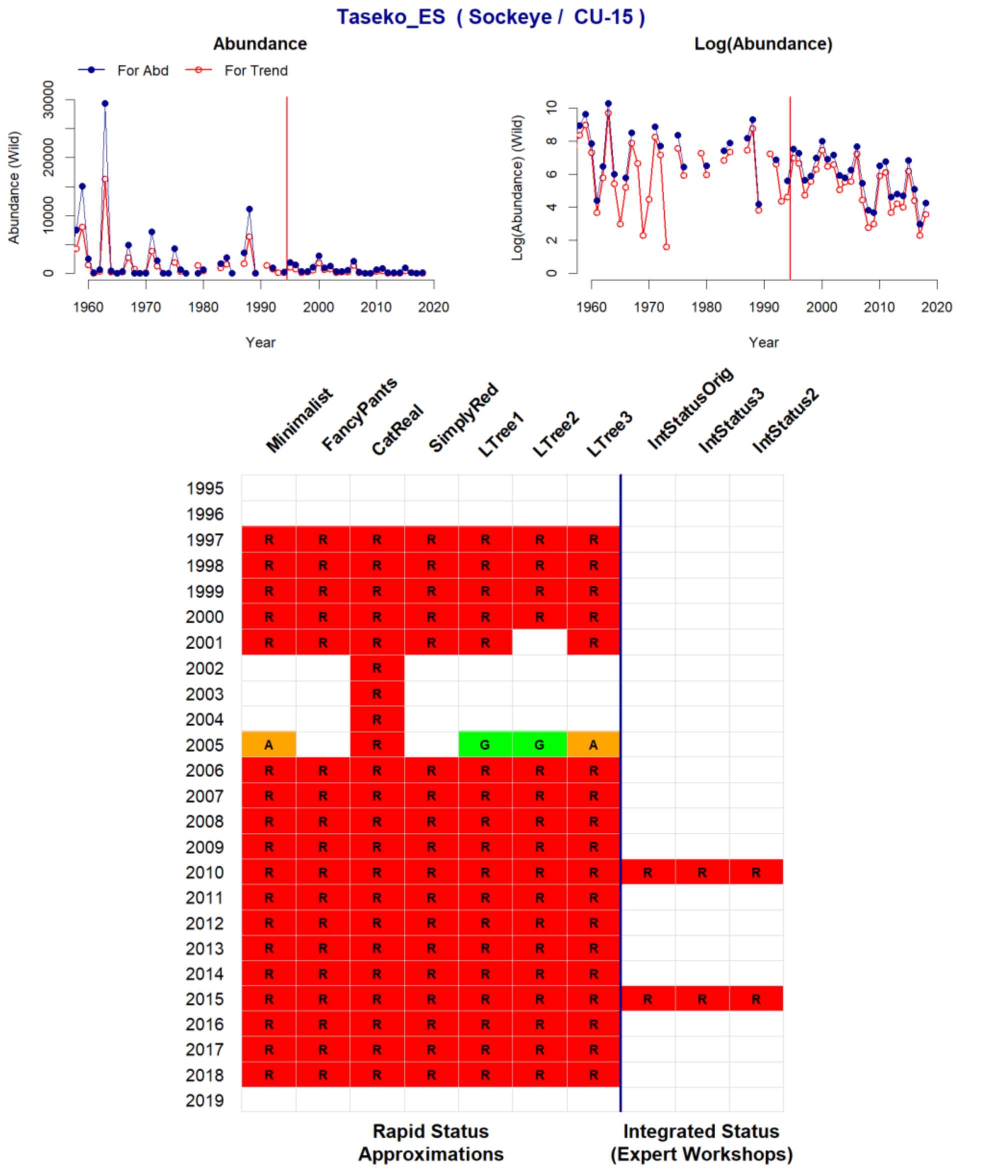


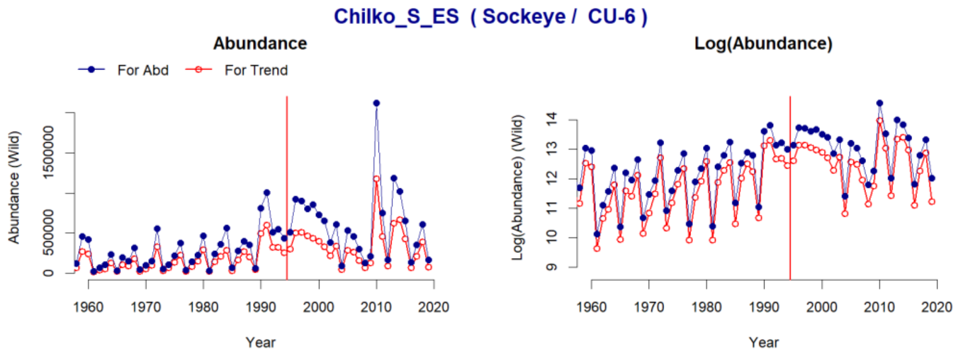
Figure 47. Retrospective test of rapid status - Fraser Sockeye – Taseko-ES. Figure shows annual rapid statuses assigned by seven alternative candidate algorithms, as well as completed integrated status assessments. Note that integrated status assessments are mapped onto the year of data used, not the year the WSP status assessment was done. Statuses are denoted as R = *Red*, RA = *Red/Amber*, A = *Amber*, AG = *Amber/Green*, G = *Green*, NR = *NotRed*, DD = *Data Deficient*, and UD = *Undetermined*. For comparison, integrated statuses are shown as the original status assignment from the expert workshops (IntStatusOrig), as well as the conversions to the 3-status (*Red/Amber/Green*) and 2-status scales (*Red/NotRed*).

H.5 FRASER SOCKEYE – SUMMER

Full integrated status assessments of Summer run sockeye were completed in 2011 (Grant and Pestal 2012) , using spawner data up to 2010, and in 2017 (Grant et al. 2020), using spawner data up to 2015. The integrated assessment and retrospective test cover 4 CUs in the Summer run sockeye management unit:

- *Chilko-S-ES* (**Figure 48**): The integrated status assessment was *Green* with data up to 2010 and *Green* with data up to 2015. All seven algorithms could assign statuses for every year since 1995. Six of the seven algorithms picked up a worsening status in the early 2000s, followed by an improvement in the 2010s. All seven algorithms indicated a worse status (*Red*, *Red/Amber*, or *Amber*) than the expert workshop (*Green*) with data up to 2010, but most of the rapid statuses switched to *Green* within a year or two. Most algorithms matched the *Green* integrated status assessment for data up to 2015. Four of the algorithms then indicate worsening status starting in 2017/2018, shortly after the last integrated status assessment was completed in 2017 with data up to 2015. For data up to to 2010, *Learning Tree 3* assigns *Amber*, one full status category worse than the integrated status assigned in the expert workshop. For data up to to 2015, *Learning Tree 3* assigns *Green* and matches the integrated status assigned in the expert workshop.
- *Francois-Fraser-S* (**Figure 49**): The integrated status assessment was *Red/Amber* with data up to 2010 and *Amber/Green* with data up to 2015. This CU accounts for 2 of 5 cases where *Learning Tree 3* assigns a better status than the expert workshop consensus, and does so with high confidence. However, the differences are actually small in terms of the original scoring, where they account for only a half-step (*Amber* vs. *Red/Amber*, *Green* vs. *Amber/Green*). The case narratives from the workshop explain that status was downgraded due to high uncertainty in the estimated benchmarks for the *relative abundance* metric (S_{gen}, S_{msy}). Section 4.4 discusses the details. The majority of algorithms, including *Learning Tree 3*, indicate *Amber* status for most years since 2006. For data up to 2010, *Learning Tree 3* assigns *Amber*, which almost matches the *Red/Amber* integrated status assigned in the expert workshop. For data up to 2015, *Learning Tree 3* assigns *Green* which almost matches the integrated status of *Amber/Green* assigned in the expert workshop.
- *Quesnel-S* (**Figure 50**): The integrated status assessment was *Red/Amber* with data up to 2010 and *Red/Amber* with data up to 2015. Six of the seven algorithms indicate *Green* or *NotRed* status from the mid-1990s to the mid-2000s, followed by worsening status, and turning into *Red* sometime around 2009. One algorithm, the *Minimalist*, points to an improvement in status around 2017/2018, but the other six algorithms continue to show poor status. For data up to 2010 and up to 2015, *Learning Tree 3* assigns *Red* status, which almost matches the integrated status of *Red/Amber* assigned in the expert workshop.
- *Takla-Trembleur-Stuart-S* (**Figure 51**): The integrated status assessment was *Red/Amber* with data up to 2010 and *Red/Amber* with data up to 2015. This CU accounts for 1 of 5 cases where *Learning Tree 3* assigns a better status than the expert workshop consensus and does so with high confidence. However, the difference is actually small in terms of the original scoring, where it accounts for only a half-step (*Amber* vs. *Red/Amber*). The case narrative from the workshop explains that status was down-graded due to high uncertainty in the estimated benchmarks for the *relative abundance* metric (S_{gen}, S_{msy}), combined with a steep decline in abundance

(i.e. *percent change*). Section 4.4 discusses the details. Most algorithms indicate *Amber* status for most years since 2006, which is a half-step above the last integrated assessment with data up to 2015.:

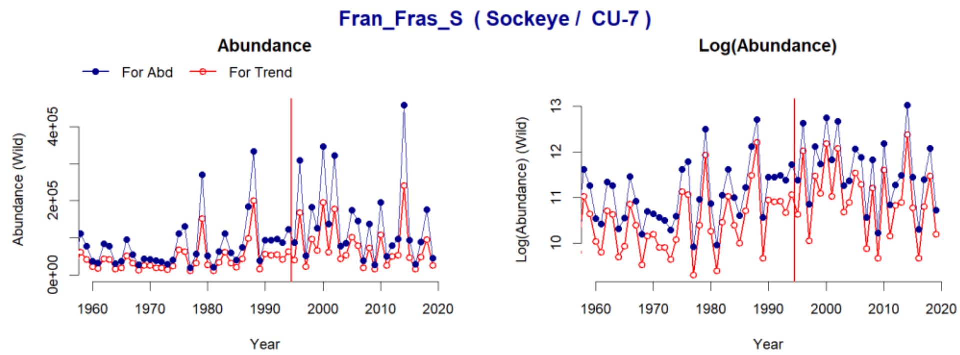


	Minimalist	FancyPants	CatReal	SimplyRed	LTree1	LTree2	LTree3	IntStatusOrig	IntStatus3	IntStatus2
1995	G	G	A	NR	G	G	G			
1996	G	G	A	NR	G	G	G			
1997	G	G	A	NR	G	G	G			
1998	G	G	A	NR	G	G	G			
1999	G	G	A	NR	G	G	G			
2000	G	G	A	NR	G	G	G			
2001	G	G	A	NR	G	G	G			
2002	G	G	A	NR	G	G	G			
2003	G	G	A	NR	G	G	G			
2004	A	G	A	NR	A	A	A			
2005	A	G	A	NR	A	A	A			
2006	A	G	A	NR	A	A	A			
2007	A	RA	A	NR	A	A	A			
2008	A	RA	A	R	A	A	A			
2009	A	RA	A	R	A	A	A			
2010	A	RA	A	R	A	A	A	G	G	NR
2011	A	RA	A	NR	G	G	G			
2012	A	G	A	NR	G	G	G			
2013	A	G	A	NR	G	G	G			
2014	A	G	A	NR	G	G	G			
2015	G	G	A	NR	G	G	G	G	G	NR
2016	G	G	A	NR	G	G	G			
2017	A	G	A	NR	G	G	G			
2018	A	G	A	NR	A	A	A			
2019	A	G	A	NR	A	A	A			

Rapid Status Approximations

Integrated Status (Expert Workshops)

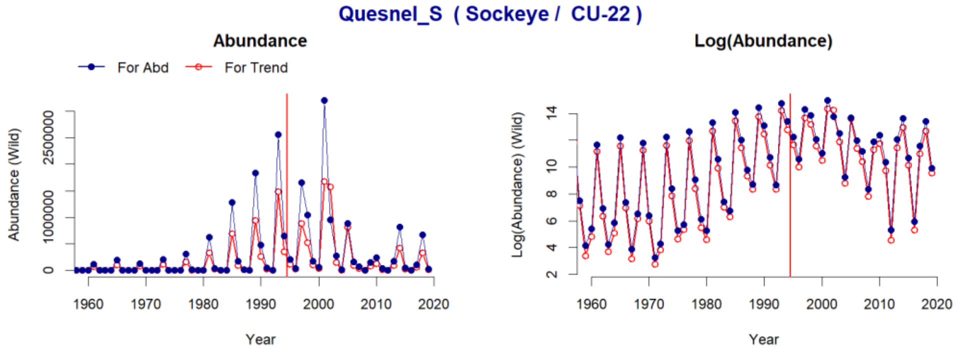
Figure 48. Retrospective test of rapid status approximations - Fraser Sockeye – Chilko-S-ES. Figure shows annual rapid statuses assigned by seven alternative candidate algorithms, as well as completed integrated status assessments. Note that integrated status assessments are mapped onto the year of data used, not the year the WSP status assessment was done. Statuses are denoted as R = *Red*, RA = *Red/Amber*, A = *Amber*, AG = *Amber/Green*, G = *Green*, NR = *NotRed*, DD = *Data Deficient*, and UD = *Undetermined*. For comparison, integrated statuses are shown as the original status assignment from the expert workshops (IntStatusOrig), as well as the conversions to the 3-status (*Red/Amber/Green*) and 2-status scales (*Red/NotRed*).



	Minimalist	FancyPants	CatReal	SimplyRed	LTree1	LTree2	LTree3	IntStatusOrig	IntStatus3	IntStatus2
1995	A	G	A	NR	A	A	A			
1996	A	G	A	NR	A	G	A			
1997	A	G	A	NR	A	A	A			
1998	A	G	A	NR	A	G	A			
1999	A	G	A	NR	G	G	G			
2000	A	G	A	NR	G	G	G			
2001	A	G	A	NR	G	G	G			
2002	A	G	A	NR	G	G	G			
2003	G	G	A	NR	G	G	G			
2004	G	G	A	NR	A	G	A			
2005	A	G	A	NR	G	G	G			
2006	A	G	A	NR	A	A	A			
2007	A	G	A	NR	A	A	A			
2008	A	G	A	NR	A	A	A			
2009	A	G	A	NR	A	A	A			
2010	A	RA	A	NR	A	A	A	RA	R	R
2011	A	RA	A	NR	A	A	A			
2012	A	RA	A	NR	A	A	A			
2013	A	RA	A	NR	A	A	A			
2014	A	G	A	NR	A	A	A			
2015	A	G	A	NR	G	G	G	AG	A	NR
2016	A	G	A	NR	A	A	A			
2017	A	G	A	NR	A	A	A			
2018	A	G	A	NR	A	A	A			
2019	A	G	A	NR	A	A	A			

Rapid Status Approximations | Integrated Status (Expert Workshops)

Figure 49. Retrospective test of rapid status approximations - Fraser Sockeye – FrancoisFraser-S. Figure shows annual rapid statuses assigned by seven alternative candidate algorithms, as well as completed integrated status assessments. Note that integrated status assessments are mapped onto the year of data used, not the year the WSP status assessment was done. Statuses are denoted as R = *Red*, RA = *Red/Amber*, A = *Amber*, AG = *Amber/Green*, G = *Green*, NR = *NotRed*, DD = *Data Deficient*, and UD = *Undetermined*. For comparison, integrated statuses are shown as the original status assignment from the expert workshops (IntStatusOrig), as well as the conversions to the 3-status (*Red/Amber/Green*) and 2-status scales (*Red/NotRed*).



	Minimalist	FancyPants	CatReal	SimplyRed	LTree1	LTree2	LTree3	IntStatusOrig	IntStatus3	IntStatus2
1995	G	G	A	NR	G	G	G			
1996	G	G	A	NR	G	G	G			
1997	G	G	A	NR	G	G	G			
1998	G	G	A	NR	G	G	G			
1999	G	G	A	NR	G	G	G			
2000	G	G	A	NR	G	G	G			
2001	G	G	A	NR	G	G	G			
2002	G	G	A	NR	G	G	G			
2003	G	G	A	NR	G	G	G			
2004	G	G	A	NR	G	G	G			
2005	G	G	A	NR	A	A	A			
2006	G	G	A	NR	A	A	A			
2007	G	G	A	NR	A	A	A			
2008	G	RA	A	NR	A	A	A			
2009	R	RA	A	R	R	R	R			
2010	R	RA	A	R	R	R	R	RA	R	R
2011	R	RA	A	R	R	R	R			
2012	R	R	A	R	R	R	R			
2013	R	A	A	R	R	R	R			
2014	R	A	A	R	R	R	R			
2015	R	RA	A	R	R	R	R	RA	R	R
2016	R	RA	A	R	R	R	R			
2017	G	RA	A	R	R	R	R			
2018	G	RA	A	R	R	R	R			
2019	A	R	A	R	R	R	R			

Rapid Status Approximations

Integrated Status (Expert Workshops)

Figure 50. Retrospective test of rapid status approximations - Fraser Sockeye – Quesnel-S. Figure shows annual rapid statuses assigned by seven alternative candidate algorithms, as well as completed integrated status assessments. Note that integrated status assessments are mapped onto the year of data used, not the year the WSP status assessment was done. Statuses are denoted as R = *Red*, RA = *Red/Amber*, A = *Amber*, AG = *Amber/Green*, G = *Green*, NR = *NotRed*, DD = *Data Deficient*, and UD = *Undetermined*. For comparison, integrated statuses are shown as the original status assignment from the expert workshops (IntStatusOrig), as well as the conversions to the 3-status (*Red/Amber/Green*) and 2-status scales (*Red/NotRed*).

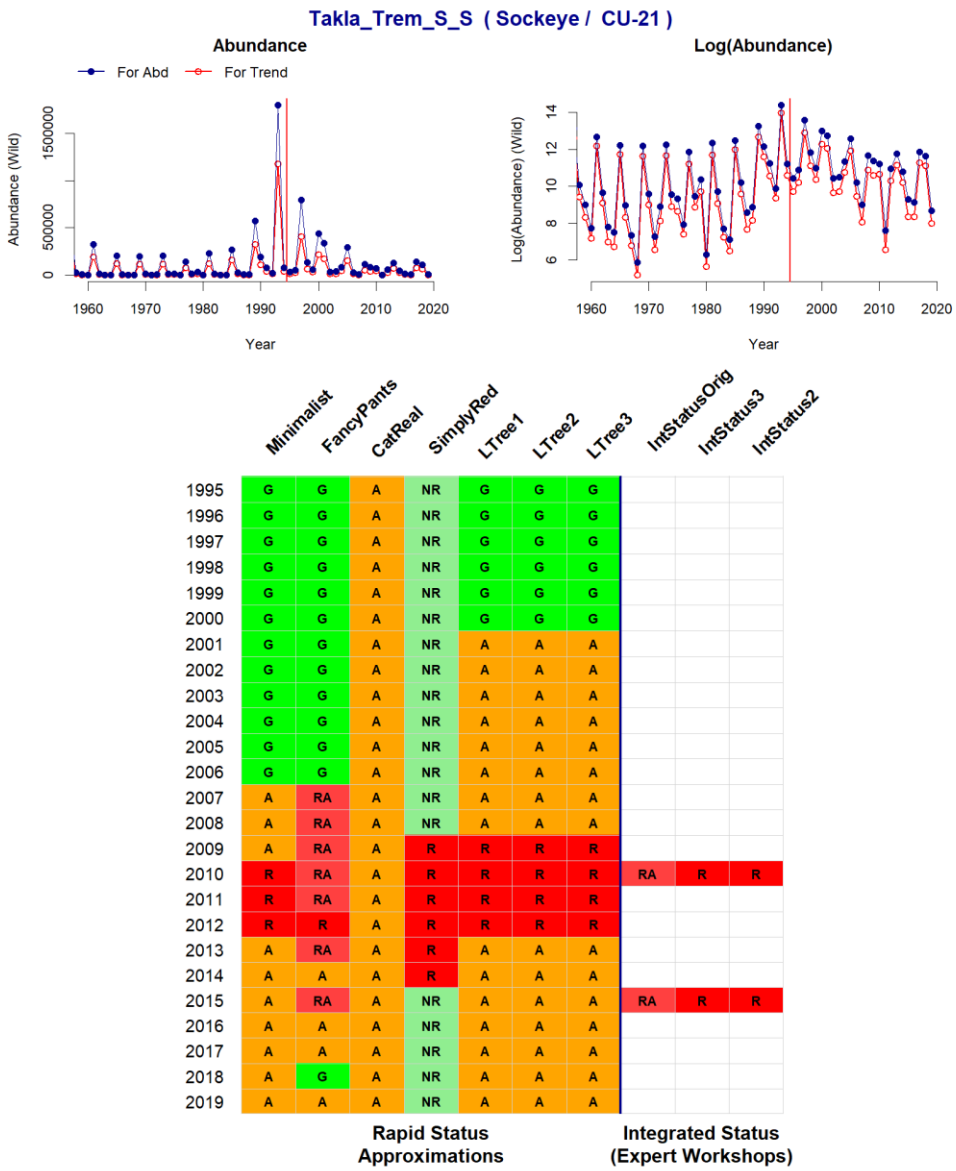


Figure 51. Retrospective test of rapid status approximations - Fraser Sockeye – Takla/Trembleur/Stuart-S. Figure shows annual rapid statuses assigned by seven alternative candidate algorithms, as well as completed integrated status assessments. Note that integrated status assessments are mapped onto the year of data used, not the year the WSP status assessment was done. Statuses are denoted as R = *Red*, RA = *Red/Amber*, A = *Amber*, AG = *Amber/Green*, G = *Green*, NR = *NotRed*, DD = *Data Deficient*, and UD = *Undetermined*. For comparison, integrated statuses are shown as the original status assignment from the expert workshops (IntStatusOrig), as well as the conversions to the 3-status (*Red/Amber/Green*) and 2-status scales (*Red/NotRed*).

H.6 FRASER SOCKEYE – LATE

Full integrated status assessments of Summer run sockeye were completed in 2011 (Grant and Pestal 2012) , using spawner data up to 2010, and in 2017 (Grant et al. 2020), using spawner data up to 2015. The integrated assessment and retrospective test covers 4 CUs in the Late run sockeye management unit:

- *Harrison-DS-L* (**Figure 52**): The integrated status assessment was *Green* with data up to 2010 and *Amber/Green* with data up to 2015. Five of the seven algorithms could assign statuses for every year since 1995. *Simply Red* could only assign status for 2 years in the retrospective, and *Fancy Pants* couldn't assign any statuses. All three versions of the *Learning Tree* algorithm picked up a worsening status in the early 2010s (from *Green* to *Amber* or *Red*), but two of the three indicate a status improvement within 2 years. Overall, the algorithms indicate no major change in status since the expert workshop assigned an integrated status of *Amber/Green* with data up to 2015. For data up to 2010, *Learning Tree 3* assigns *Green* status and matches the integrated status assigned in the expert workshop. For data up to 2015, *Learning Tree 3* assigns *Red* status, much worse than the *Amber/Green* integrated status assigned by experts. *Learning Tree 3* assigned *Red* due to the steep decline (*Percent change* = -71%), but workshop participants down-weighted this metric because the decline came after very large spawner abundances in the early 2000s.
- *Harrison-US-L* (**Figure 53**): The integrated status assessment was *Amber* with data up to 2010 and *Red* with data up to 2015. All seven algorithms could assign statuses for every year since 1995. All three versions of the *Learning Tree* algorithm match the expert assessments, assigning *Amber* status for 1995 to 2014, then switching to *Red* in 2015. Overall, all the algorithms indicate that status has not improved since the last integrated expert assessment.
- *Lillooet-Harrison-L* (**Figure 54**): The integrated status assessment was *Green* with data up to 2010 and *Amber* with data up to 2015. All seven algorithms could assign statuses for every year since 1995. Five of the seven algorithms indicated a worse status than the expert workshop for 2010 (*Amber* vs. *Green*), but by 2015 the algorithms and experts assign the same *Amber* status. Six of the seven algorithms indicate a worsening status since the last integrated assessment, switching to *Red* in either 2017 or 2019. *Learning Tree 3* assigns *Amber* for data up to 2010, one full status category worse than the expert assessment. For data up to 2015, *Learning Tree 3* assigns *Amber* and matches the expert assessment.
- *Shuswap-L* (**Figure 55**): The integrated status assessment was *Green* with data up to 2010 and *Amber/Green* with data up to 2015. All seven algorithms could assign statuses for every year since 1995. Rapid statuses for this CU differ a lot between algorithms, with the three fitted algorithms and the *Simply Red* algorithm frequently assigning worse statuses than the three versions of the *Learning Tree* algorithm. The *Learning Tree* algorithms generate statuses that closely match the expert assessments with data up to 2010 and up to 2015, and all the *Learning Tree* algorithms indicate a worsening status after the last expert assessment, shifting from *Green* to *Amber* in 2017 or 2018.

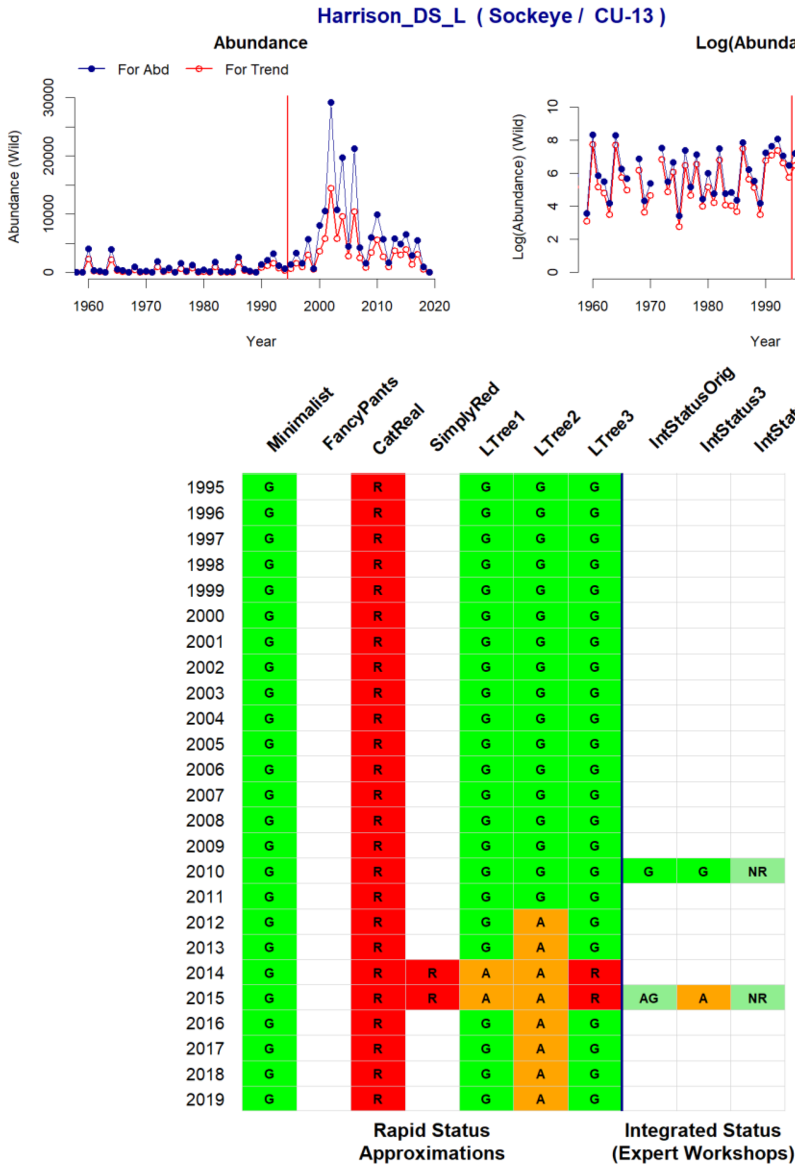
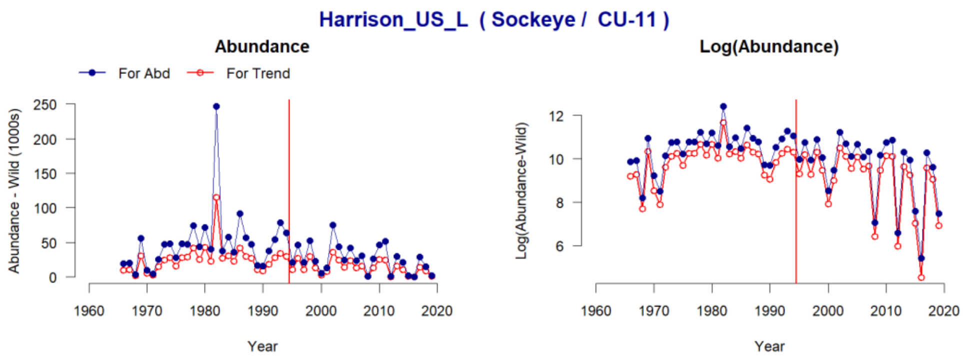


Figure 52. Retrospective test of rapid status approximations - Fraser Sockeye – Harrison-DS-L. Figure shows annual rapid statuses assigned by seven alternative candidate algorithms, as well as completed integrated status assessments. Note that integrated status assessments are mapped onto the year of data used, not the year the WSP status assessment was done. Statuses are denoted as R = *Red*, RA = *Red/Amber*, A = *Amber*, AG = *Amber/Green*, G = *Green*, NR = *NotRed*, DD = *Data Deficient*, and UD = *Undetermined*. For comparison, integrated statuses are shown as the original status assignment from the expert workshops (IntStatusOrig), as well as the conversions to the 3-status (*Red/Amber/Green*) and 2-status scales (*Red/NotRed*).

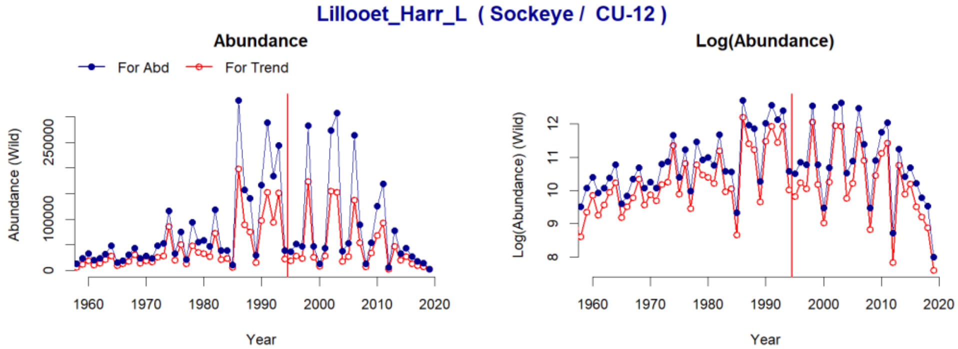


	Minimalist	FancyPants	CatReal	SimplyRed	LTree1	LTree2	LTree3	IntStatusOrig	IntStatus3	IntStatus2
1995	A	G	A	NR	A	A	A			
1996	A	G	A	NR	A	A	A			
1997	A	G	A	NR	A	A	A			
1998	A	G	A	NR	A	A	A			
1999	A	G	A	NR	A	A	A			
2000	R	R	A	R	A	A	A			
2001	R	R	A	R	A	A	A			
2002	R	R	A	R	A	A	A			
2003	R	R	A	R	A	A	A			
2004	R	R	A	R	A	A	A			
2005	A	G	A	NR	A	A	A			
2006	A	G	A	NR	A	A	A			
2007	A	A	A	NR	A	A	A			
2008	A	A	A	NR	A	A	A			
2009	R	R	A	R	A	A	A			
2010	R	R	A	R	A	A	A	A	A	NR
2011	R	R	A	R	A	A	A			
2012	R	R	A	R	A	A	A			
2013	R	R	A	R	A	A	A			
2014	R	R	A	R	A	A	A			
2015	R	R	A	R	R	R	R	R	R	R
2016	R	R	A	R	R	R	R			
2017	R	R	A	R	R	R	R			
2018	R	R	A	R	R	R	R			
2019	R	R	A	R	R	R	R			

Rapid Status Approximations

Integrated Status (Expert Workshops)

Figure 53. Retrospective test of rapid status approximations - Fraser Sockeye – Harrison-US-L. Figure shows annual rapid statuses assigned by seven alternative candidate algorithms, as well as completed integrated status assessments. Note that integrated status assessments are mapped onto the year of data used, not the year the WSP status assessment was done. Statuses are denoted as R = *Red*, RA = *Red/Amber*, A = *Amber*, AG = *Amber/Green*, G = *Green*, NR = *NotRed*, DD = *Data Deficient*, and UD = *Undetermined*. For comparison, integrated statuses are shown as the original status assignment from the expert workshops (IntStatusOrig), as well as the conversions to the 3-status (*Red/Amber/Green*) and 2-status scales (*Red/NotRed*).

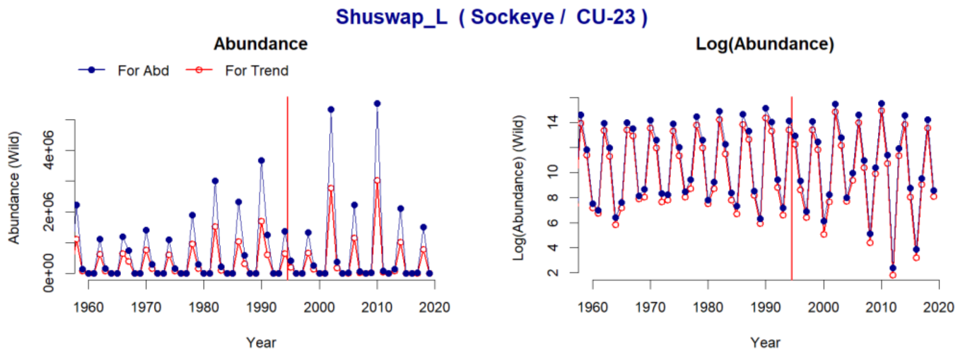


	Minimalist	FancyPants	CatReal	SimplyRed	LTree1	LTree2	LTree3	IntStatusOrig	IntStatus3	IntStatus2
1995	G	G	A	NR	G	G	G			
1996	G	G	A	NR	A	A	A			
1997	A	G	A	NR	A	A	A			
1998	A	G	A	NR	A	A	A			
1999	A	G	A	NR	A	A	A			
2000	A	RA	A	NR	A	A	A			
2001	A	RA	A	NR	A	A	A			
2002	A	RA	A	NR	A	A	A			
2003	A	RA	A	NR	A	G	A			
2004	A	G	A	NR	G	G	G			
2005	A	G	A	NR	G	G	G			
2006	A	G	A	NR	G	G	G			
2007	A	G	A	NR	A	G	A			
2008	A	G	A	NR	A	A	A			
2009	A	G	A	NR	A	A	A			
2010	A	G	A	NR	A	A	A	G	G	NR
2011	A	G	A	NR	A	A	A			
2012	A	G	A	NR	A	A	A			
2013	A	G	A	NR	A	A	A			
2014	A	RA	A	NR	A	A	A			
2015	A	A	A	NR	A	A	A	A	A	NR
2016	A	RA	A	NR	A	A	A			
2017	R	R	A	R	A	A	A			
2018	R	R	A	R	A	A	A			
2019	R	R	A	R	R	R	R			

Rapid Status Approximations

Integrated Status (Expert Workshops)

Figure 54. Retrospective test of rapid status approximations - Fraser Sockeye – Lillooet-Harrison-L. Figure shows annual rapid statuses assigned by seven alternative candidate algorithms, as well as completed integrated status assessments. Note that integrated status assessments are mapped onto the year of data used, not the year the WSP status assessment was done. Statuses are denoted as R = *Red*, RA = *Red/Amber*, A = *Amber*, AG = *Amber/Green*, G = *Green*, NR = *NotRed*, DD = *Data Deficient*, and UD = *Undetermined*. For comparison, integrated statuses are shown as the original status assignment from the expert workshops (IntStatusOrig), as well as the conversions to the 3-status (*Red/Amber/Green*) and 2-status scales (*Red/NotRed*).



Year	Minimalist	FancyPants	CatReal	SimplyRed	LTree1	LTree2	LTree3	IntStatusOrig	IntStatus3	IntStatus2
1995	A	G	A	NR	A	A	A			
1996	A	G	A	NR	A	A	A			
1997	A	G	A	NR	A	A	A			
1998	A	G	A	NR	A	A	A			
1999	A	G	A	NR	A	A	A			
2000	R	R	A	R	A	A	A			
2001	R	R	A	R	A	A	A			
2002	R	R	A	R	G	G	G			
2003	R	R	A	R	G	G	G			
2004	A	G	A	NR	G	G	G			
2005	A	G	A	NR	G	G	G			
2006	A	G	A	NR	G	G	G			
2007	A	G	A	NR	G	G	G			
2008	A	A	A	NR	G	G	G			
2009	A	A	A	NR	G	G	G			
2010	R	R	A	R	G	G	G	G	G	NR
2011	R	R	A	R	G	G	G			
2012	R	R	A	R	G	G	G			
2013	R	R	A	R	G	G	G			
2014	R	R	A	R	G	G	G			
2015	R	R	A	R	G	G	G	AG	A	NR
2016	R	R	A	R	G	G	G			
2017	R	R	A	R	G	G	A			
2018	R	R	A	R	A	A	A			
2019	R	R	A	R	A	A	A			

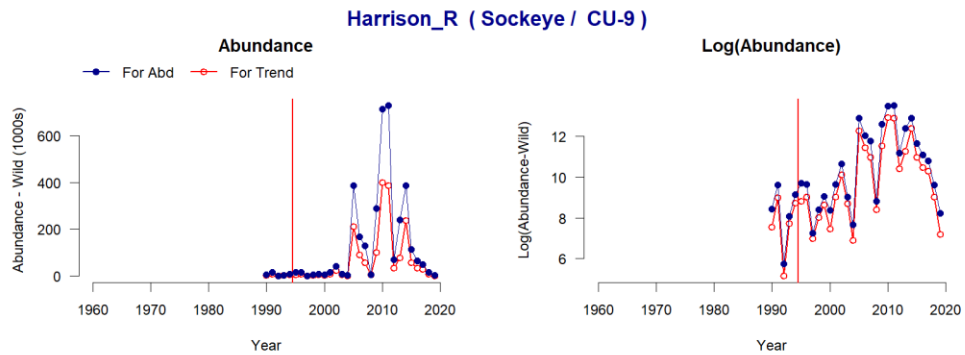
Rapid Status Approximations — Integrated Status (Expert Workshops)

Figure 55. Retrospective test of rapid status approximations - Fraser Sockeye – Harrison-DS-L. Figure shows annual rapid statuses assigned by seven alternative candidate algorithms, as well as completed integrated status assessments. Note that integrated status assessments are mapped onto the year of data used, not the year the WSP status assessment was done. Statuses are denoted as R = *Red*, RA = *Red/Amber*, A = *Amber*, AG = *Amber/Green*, G = *Green*, NR = *NotRed*, DD = *Data Deficient*, and UD = *Undetermined*. For comparison, integrated statuses are shown as the original status assignment from the expert workshops (IntStatusOrig), as well as the conversions to the 3-status (*Red/Amber/Green*) and 2-status scales (*Red/NotRed*).

H.7 FRASER SOCKEYE – RIVER-TYPE

Full integrated status assessments of Summer run sockeye were completed in 2011 (Grant and Pestal 2012) , using spawner data up to 2010, and in 2017 (Grant et al. 2020), using spawner data up to 2015. The integrated assessment and retrospective test covers 4 CUs in the Late run sockeye management unit:

- *Harrison-R* (**Figure 56**): The integrated status assessment was *Green* with data up to 2010 and *Green* with data up to 2015. All seven algorithms could assign statuses for every year since 2007, and four algorithms could assign status since 1995. The three *Learning Tree* algorithms assigned *Red* status for the early part of the retrospective, from 1995 to 2006, then improving gradually (e.g. switched from *Red* to *Amber* in 2007, then to *Green* in 2011, one year after the expert workshop assigned and integrated status of *Green*). With data up to 2015, six of the seven algorithms match the expert assessment of *Amber*. Overall, five of the seven algorithms indicate a worsening status since the last expert assessment. The *Learning Tree* algorithms switch to *Amber* in 2017, and to *Red* in 2019. *Minimalist*, *Fancy Pants* and *Simply Red* all indicate a stable *Green* or *NotRed* for 2007 to 2018, then a switch to a poorer status: *Amber* for *Minimalist*, *Red* for *Fancy Pants* and *Simply Red.*
- *Widgeon-RT* (**Figure 57**): The integrated status assessment was *Red* with data up to 2010 and *Red* with data up to 2015. Only *Minimalist* and the three versions of the *Learning Tree* algorithm could assign statuses for all years since 1997. Most of the completed rapid status assignments since 1995 are *Red*. Two of the algorithms indicated an improvement in status in the early 2010s, but by 2014 all three *Learning Tree* versions assigned a *Red* status, and by 2018 all six algorithms that could assign a rapid status assign *Red*. Overall, there is no indication that status has improved since the last expert assessment with data up to 2015.



	Minimalist	FancyPants	CatReal	SimplyRed	LTree1	LTree2	LTree3	IntStatusOrig	IntStatus3	IntStatus2
1995			A		R	R	R			
1996			A		R	R	R			
1997			A		R	R	R			
1998			A		R	R	R			
1999			A		R	R	R			
2000			A		R	R	R			
2001			A		R	R	R			
2002			A		R	R	R			
2003			A		R	R	R			
2004			A		R	R	R			
2005			A		R	R	R			
2006			A		R	R	R			
2007	G	G	A	NR	A	A	A			
2008	G	G	A	NR	A	A	A			
2009	G	G	A	NR	A	A	A			
2010	G	G	A	NR	A	A	A	G	G	NR
2011	G	G	A	NR	G	G	G			
2012	G	G	A	NR	G	G	G			
2013	G	G	A	NR	G	G	G			
2014	G	G	A	NR	G	G	G			
2015	G	G	A	NR	G	G	G	G	G	NR
2016	G	G	A	NR	G	G	G			
2017	G	G	A	NR	A	A	A			
2018	G	G	A	NR	A	A	A			
2019	A	R	A	R	R	R	R			

Rapid Status Approximations | Integrated Status (Expert Workshops)

Figure 56. Retrospective test of rapid status approximations - Fraser Sockeye – Harrison-R. Figure shows annual rapid statuses assigned by seven alternative candidate algorithms, as well as completed integrated status assessments. Note that integrated status assessments are mapped onto the year of data used, not the year the WSP status assessment was done. Statuses are denoted as R = *Red*, RA = *Red/Amber*, A = *Amber*, AG = *Amber/Green*, G = *Green*, NR = *NotRed*, DD = *Data Deficient*, and UD = *Undetermined*. For comparison, integrated statuses are shown as the original status assignment from the expert workshops (IntStatusOrig), as well as the conversions to the 3-status (*Red/Amber/Green*) and 2-status scales (*Red/NotRed*).

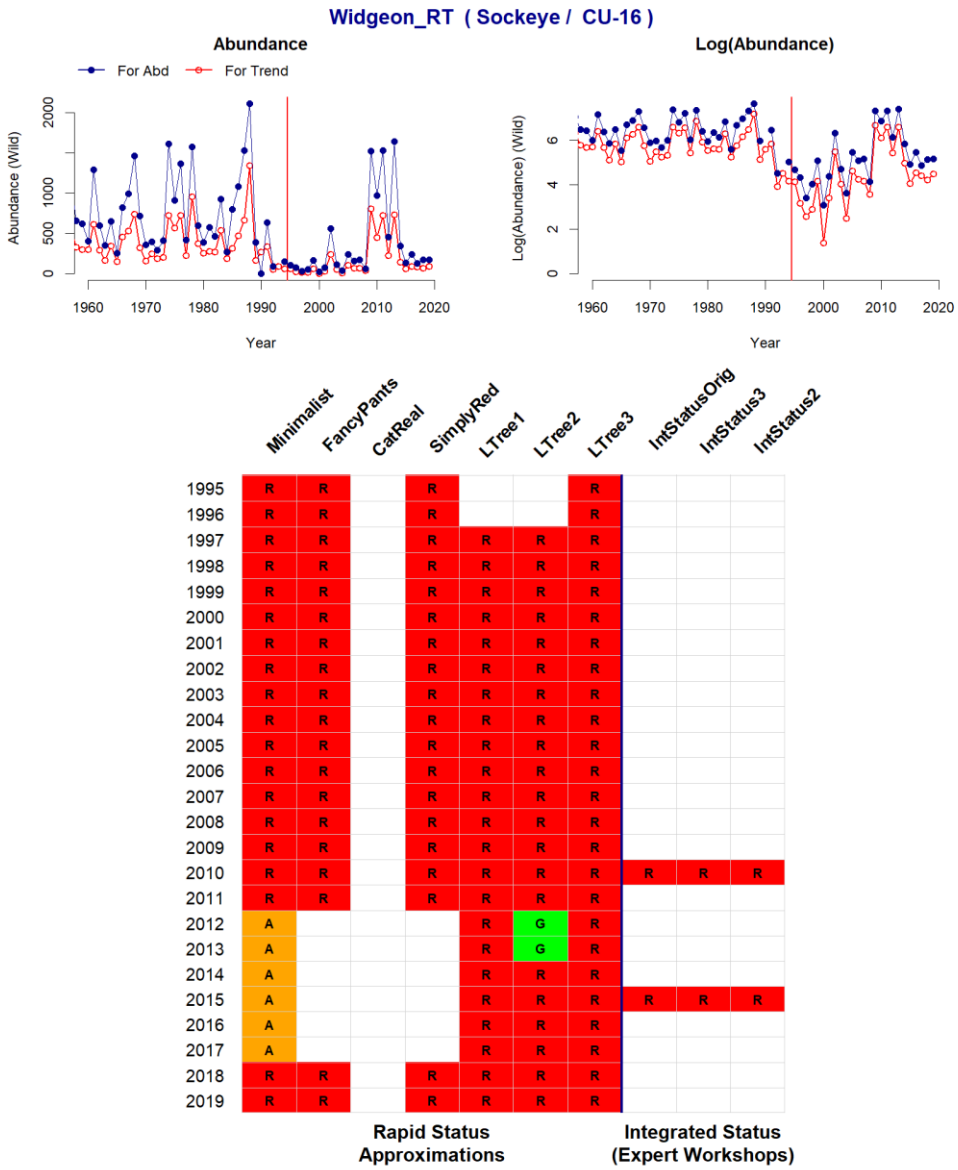


Figure 57. Retrospective test of rapid status approximations - Fraser Sockeye – Widgeon-RT. Figure shows annual rapid statuses assigned by seven alternative candidate algorithms, as well as completed integrated status assessments. Note that integrated status assessments are mapped onto the year of data used, not the year the WSP status assessment was done. Statuses are denoted as R = *Red*, RA = *Red/Amber*, A = *Amber*, AG = *Amber/Green*, G = *Green*, NR = *NotRed*, DD = *Data Deficient*, and UD = *Undetermined*. For comparison, integrated statuses are shown as the original status assignment from the expert workshops (IntStatusOrig), as well as the conversions to the 3-status (*Red/Amber/Green*) and 2-status scales (*Red/NotRed*).

H.8 SOUTHERN BC CHINOOK – FRASER - LOWER

Full integrated status assessments of Lower Fraser Chinook were completed in 2012 (DFO 2016) , using spawner data up to 2012. The integrated assessment and retrospective test covers 5 CUs in the Lower Fraser Chinook management unit:

- *Lower Fraser River_FA_0.3* (**Figure 58**): The integrated status assessment was *Green* with data up to 2012. All seven algorithms could assign statuses for every year since 1995. The *Categorical Realist* assigned *Amber* status for all years. The other algorithms all assign changing status over time, with worse statuses up to 1998, better statuses for 1999 to 2007, worse statuses for 2008 to 2010/2011, brief improvement around 2012, and then worse status. Six of the seven algorithms assign *Red* status for 2019. Five of the algorithms match the expert assessment for 2012, while the other two assign a worse status.
- *Lower Fraser River_SP_1.3* (**Figure 59**): The integrated status assessment was *To Be Determined* with data up to 2012. At the time, there was no site classified as wild, and therefore the CU was not assessed. However, experts in the workshop noted that "*the classification of enhancement level needs to be reviewed because enhancement stopped in 2002 brood year and the system now has natural spawners. There are also a number of locations within this TU that have no enhancement but are not surveyed."* Since then, the site classifications have been updated (Brown et al. 2020). All 7 algorithms could assign statuses since 2017 and all of them assign *Red*. Before 2017 4 of the algorithms could assign statuses back to 2000, and 1 algorithm back to 1995. For 2016 and earlier the statuses assigned by the algorithms ranging from *Red* to *Green*.
- *Lower Fraser River-Upper Pitt_SU_1.3* (No Figure): Notes from the workshop in 2012 state: Based on available data and the metrics presented, most groups assessed this CU as *Red* due to declining trends and low abundance. However, participants agreed to a DD assessment based on additional information provided by a participating expert (the single site with data is not representative, and surveys of additional sites within the CU are currently not feasible). Specifically, the rationale was "Time series of good quality data available, but considered not representative of whole CU. Only 1 population surveyed but others may exist that are not yet known." Therefore, the CU was not assessed in either the expert workshop or in the retrospective test.
- *Lower Fraser River_SU_1.3* (No Figure): Notes from the workshop in 2012 state: Time series of good quality data available but considered not representative of whole CU. Data available for only 1 site out of 7 (most abundant site cannot be assessed due to low visibility), and for the site with data, the time series is too short. Therefore, the CU was not assessed in either the expert workshop or in the retrospective test.
- *Maria Slough_SU_0.3* (No Figure): Notes from the workshop in 2012 state: The CU has received an enormous amount of stewardship and watershed restoration activity. Human land-use impacts have changed the hydrography of this geographically small CU. There is no data for wild sites in the CU. Therefore, the CU was not assessed in either the expert workshop or in the retrospective test.

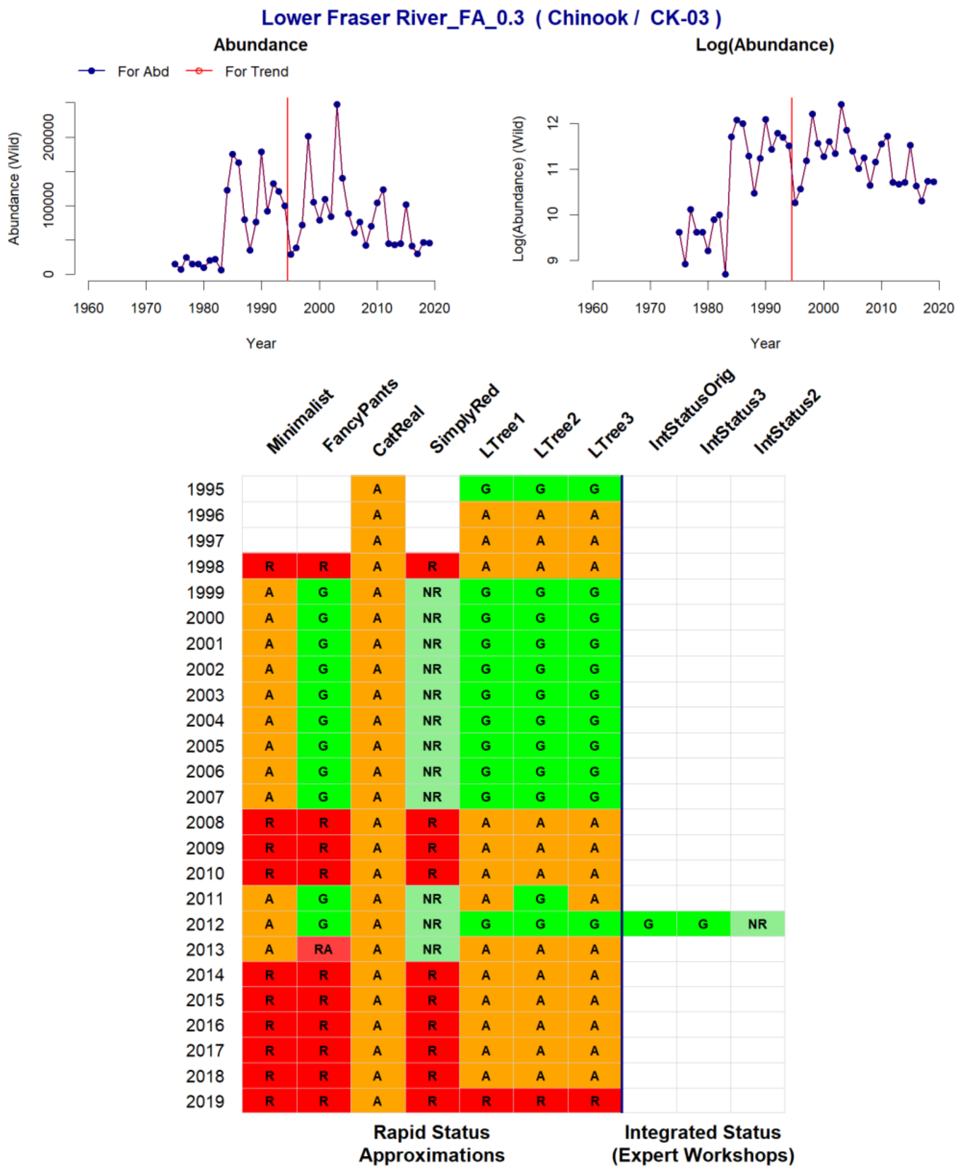


Figure 58. Retrospective test of rapid status approximations - Fraser Sockeye – Lower Fraser River_FA_1.3. Figure shows annual rapid statuses assigned by seven alternative candidate algorithms, as well as completed integrated status assessments. Note that integrated status assessments are mapped onto the year of data used, not the year the WSP status assessment was done. Statuses are denoted as R = *Red*, RA = *Red/Amber*, A = *Amber*, AG = *Amber/Green*, G = *Green*, NR = *NotRed*, DD = *Data Deficient*, and UD = *Undetermined*. For comparison, integrated statuses are shown as the original status assignment from the expert workshops (IntStatusOrig), as well as the conversions to the 3-status (*Red/Amber/Green*) and 2-status scales (*Red/NotRed*).

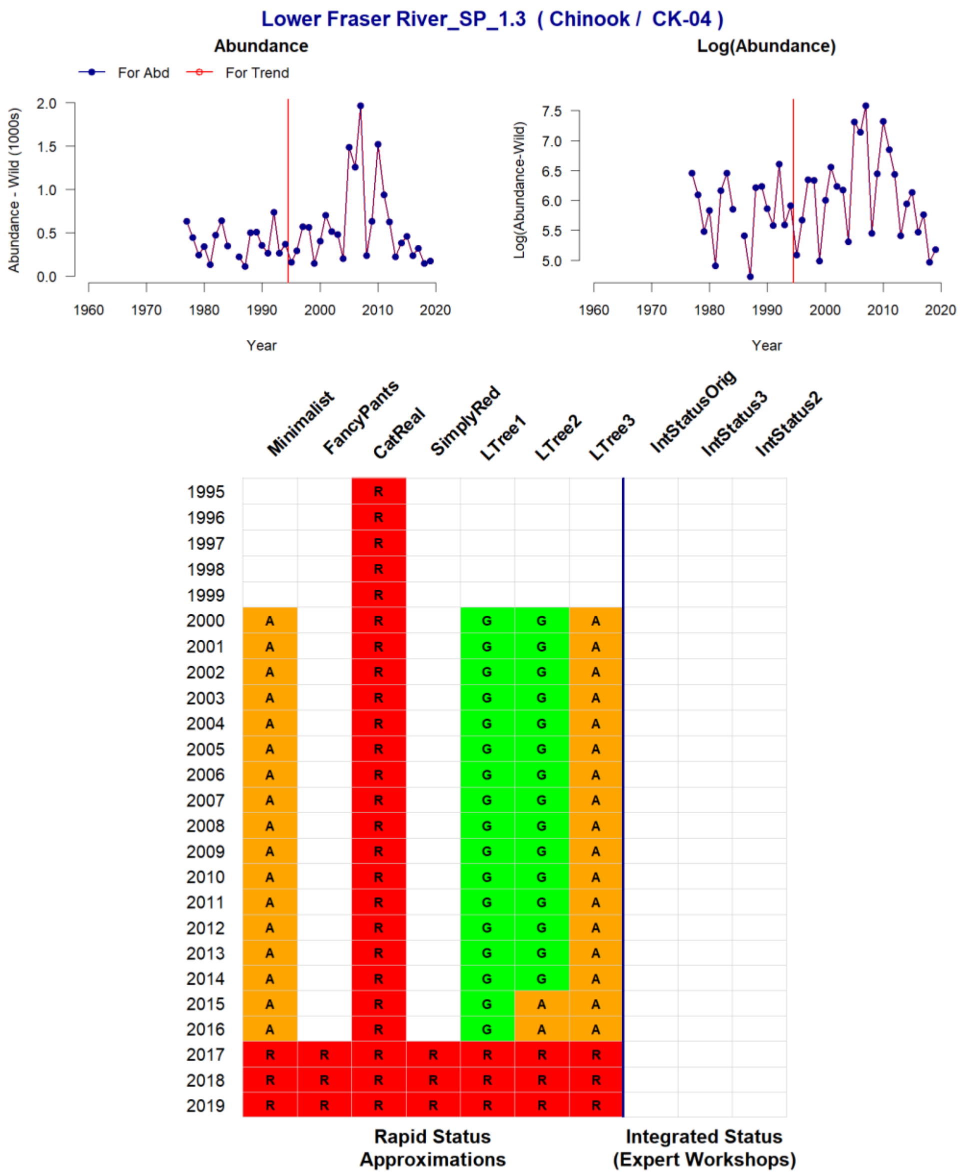


Figure 59. Retrospective test of rapid status approximations – SBC Chinook – Lower Fraser River_SP_1.3. Figure shows annual rapid statuses assigned by seven alternative candidate algorithms, as well as completed integrated status assessments. Note that integrated status assessments are mapped onto the year of data used, not the year the WSP status assessment was done. Statuses are denoted as R = *Red*, RA = *Red/Amber*, A = *Amber*, AG = *Amber/Green*, G = *Green*, NR = *NotRed*, DD = *Data Deficient*, and UD = *Undetermined*. For comparison, integrated statuses are shown as the original status assignment from the expert workshops (IntStatusOrig), as well as the conversions to the 3-status (*Red/Amber/Green*) and 2-status scales (*Red/NotRed*).

H.9 SOUTHERN BC CHINOOK – FRASER - UPPER

Full integrated status assessments of Upper Fraser Chinook were completed in 2012 (DFO 2016), using spawner data up to 2012. The integrated assessment and retrospective test covers 5 CUs in the Upper Fraser Chinook management unit:

- *Middle Fraser-Fraser Canyon_SP_1.3* (No Figure): Data available, but none meet the quality criteria. Only records are opportunistic observations during Sockeye Salmon surveys. Therefore, the CU was not assessed in either the expert workshop or in the retrospective test.
- *Middle Fraser River-Portage_FA_1.3* (**Figure 60**): The integrated status assessment was *Red* with data up to 2012. All seven algorithms could assign statuses for every year since 2014, and they all assigned *Red* status for all years since then.
- *Middle Fraser River_SP_1.3* (**Figure 61**): The integrated status assessment was *Red* with data up to 2012. All seven algorithms assigned *Red* for 2009 to 2013. From 2014 to 2017 a variable number of algorithms could assign status, and statuses differed between algorithms but stayed mostly stable for each algorithm. All seven algorithms assign *Red* status for 2019. For 2012, all seven algorithms matched the expert assessment of *Red*.
- *Middle Fraser River_SU_1.3* (No Figure): Experts in the status workshop completed a status assessment of this CU and assigned *Amber* status due to mixed signals across metrics. A subsequent review (Brown et al. 2020) found that the available data are not usable for WSP metrics, due to site-specific challenges. For example, "*the Stuart River has a large number of fish, up to 15,000 in some years; however, the percentage of the fish that are counted is unknown and varies annually depending on the water clarity. Winds on Stuart Lake disturb the shoreline sediments and can lead to visibility of less than 1m in some years, whereas in others, visibility can be up to 4m. In the mid-2000s, the noise in the time series was believed to exceed any signal and the surveys were dropped from the monitoring program*. Therefore, the CU was not assessed in either the expert workshop or in the retrospective test.
- *Upper Fraser River_SP_1.3* (Figure 62): The integrated status assessment was *Red* with data up to 2012. All seven algorithms assigned *Red* for 2009 to 2015 and matched the expert assessment for 2012. For 2016 and 2017, five algorithms could assign status, and statuses ranged from *Red* to *Green*. All seven algorithms switched back to *Red* status starting in 2018.

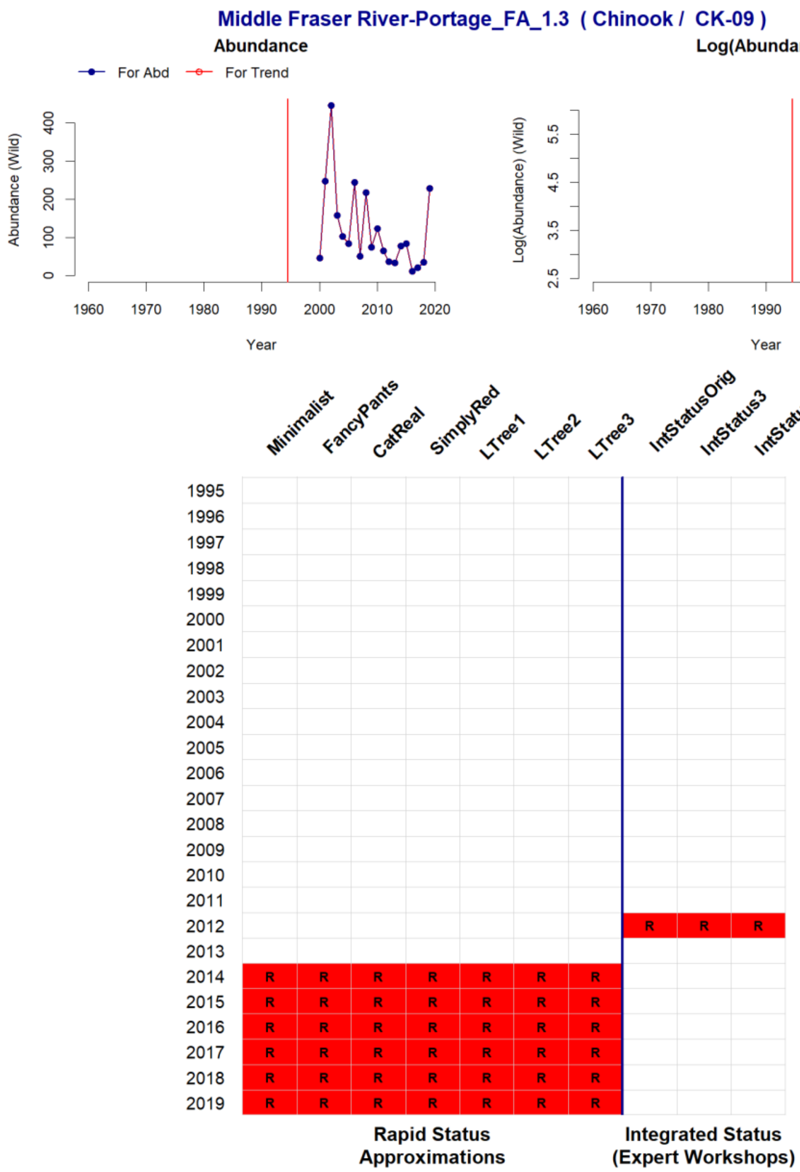


Figure 60. Retrospective test of rapid status approximations – Upper Fraser Chinook – Middle Fraser River-Portage_FA_1.3. Figure shows annual rapid statuses assigned by seven alternative candidate algorithms, as well as completed integrated status assessments. Note that integrated status assessments are mapped onto the year of data used, not the year the WSP status assessment was done. Statuses are denoted as R = *Red*, RA = *Red/Amber*, A = *Amber*, AG = *Amber/Green*, G = *Green*, NR = *NotRed*, DD = *Data Deficient*, and UD = *Undetermined*. For comparison, integrated statuses are shown as the original status assignment from the expert workshops (IntStatusOrig), as well as the conversions to the 3-status (*Red/Amber/Green*) and 2-status scales (*Red/NotRed*).

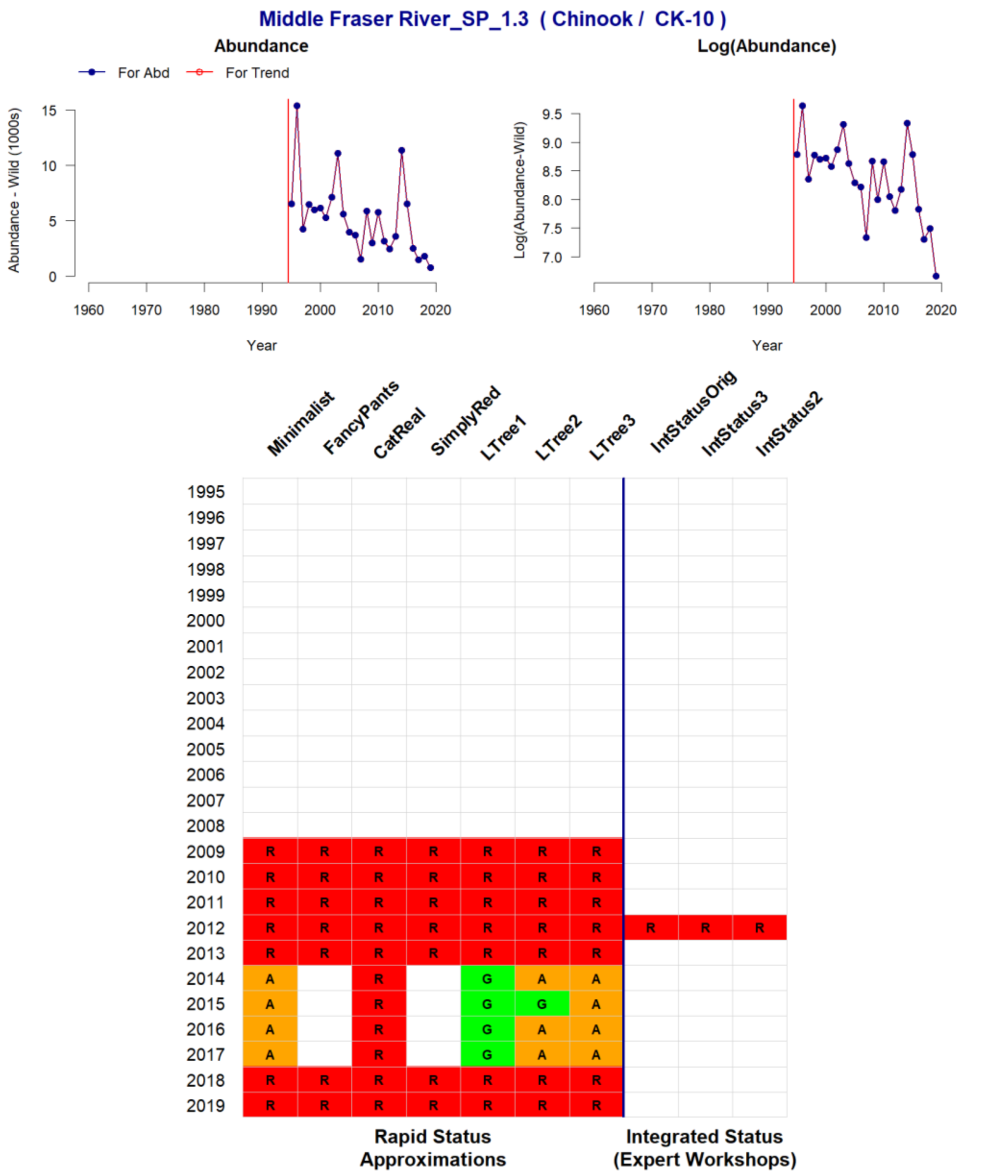
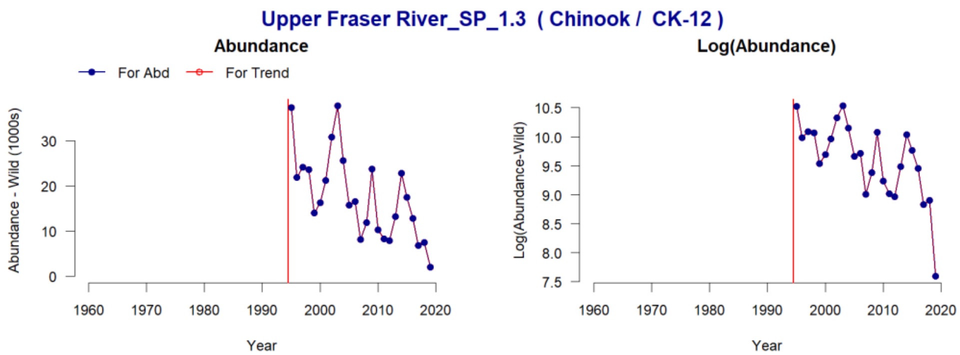


Figure 61. Retrospective test of rapid status approximations - Upper Fraser Chinook - Middle Fraser River_SP_1.3. Figure shows annual rapid statuses assigned by seven alternative candidate algorithms, as well as completed integrated status assessments. Note that integrated status assessments are mapped onto the year of data used, not the year the WSP status assessment was done. Statuses are denoted as R = *Red*, RA = *Red/Amber*, A = *Amber*, AG = *Amber/Green*, G = *Green*, NR = *NotRed*, DD = *Data Deficient*, and UD = *Undetermined*. For comparison, integrated statuses are shown as the original status assignment from the expert workshops (IntStatusOrig), as well as the conversions to the 3-status (*Red/Amber/Green*) and 2-status scales (*Red/NotRed*).



	Minimalist	FancyPants	CatReal	SimplyRed	LTree1	LTree2	LTree3	IntStatusOrig	IntStatus3	IntStatus2
1995										
1996										
1997										
1998										
1999										
2000										
2001										
2002										
2003										
2004										
2005										
2006										
2007										
2008										
2009	R	R	R	R	R	R	R			
2010	R	R	R	R	R	R	R			
2011	R	R	R	R	R	R	R			
2012	R	R	R	R	R	R	R	R	R	R
2013	R	R	R	R	R	R	R			
2014	R	R	R	R	R	R	R			
2015	R	R	R	R	R	R	R			
2016	A		R		G	A	A			
2017	A		R		G	A	A			
2018	R	R	R	R	R	R	R			
2019	R	R	R	R	R	R	R			

Figure 62. ***Retrospective test of rapid status approximations - Upper Fraser Chinook – Upper Fraser River_SP_1.3.*** Figure shows annual rapid statuses assigned by seven alternative candidate algorithms, as well as completed integrated status assessments. Note that integrated status assessments are mapped onto the year of data used, not the year the WSP status assessment was done. Statuses are denoted as R = *Red*, RA = *Red/Amber*, A = *Amber*, AG = *Amber/Green*, G = *Green*, NR = *NotRed*, DD = *Data Deficient*, and UD = *Undetermined*. For comparison, integrated statuses are shown as the original status assignment from the expert workshops (IntStatusOrig), as well as the conversions to the 3-status (*Red/Amber/Green*) and 2-status scales (*Red/NotRed*).

H.10 SOUTHERN BC CHINOOK – FRASER - THOMPSON

Full integrated status assessments of Thompson Chinook were completed in 2012 (DFO 2016), using spawner data up to 2012. The integrated assessment and retrospective test cover 8 CUs in the Upper Fraser Chinook management unit:

- *South Thompson_SU_0.3* (**Figure 63**): The integrated status assessment was *Green* with data up to 2012. Five of the algorithms, including *Learning Tree 3* could assign status since 2011, with each algorithm assigning a stable status across all years, but statuses ranging from *Red* to *Green* across algorithms. Two algorithms could not assign status for any year. *Learning Trees 1 and 2* assigned *Green* for 2012*,* matching the expert assessment, but *Learning Tree 3* assigned *Amber*.
- *South Thompson_SU_1.3* (**Figure 64**): The integrated status assessment was *Red/Amber* with data up to 2012. Five of the algorithms, including *Learning Tree 3* could assign status since 2013, with four algorithms assigning a stable status across all years, but statuses ranging from *Red* to *Green* across algorithms. Two algorithms could not assign status for any year. No algorithm could assign a rapid status for 2012, because updated data based on Brown *et al.* (2020) excludes some earlier records, and therefore trend metrics can now only be calculated starting in 2013.
- *Shuswap River_SU_0.3* (Figure 65): Experts at the status workshop (DFO 2016) designated this CU as one of 11 SBC Chinook CUs that are type-4 data deficient (i.e. (good quality data is available, but none for wild sites). Since then, the lower Shuswap river spawning area (nuSEDS popID = 46437) was reclassified as having a low level of enhanced contribution, and therefore a time series for CU status assessment is now available. Prior to 2018, four algorithms (*Minimalist* and *Learning Trees* 1-3) could assign a status every years since 1997, and *Fancy Pants* could assign status for 1999-2004. Statuses are stable over time for three of the four algorithms, but range from *Amber* to *Green* across algorithms for each year. Six algorithms could assign status for 2018-2019, and all indicate a worsening status.
- *South Thompson-Bessette Creek_SU_1.2* (No Figure): Experts at the status workshop (DFO 2016) assigned a provisional *Red* status, noting "*precipitous decline and extremely low numbers (but need to revisit CU definition). If this is accepted as a CU, then no question that the population has declined drastically."* Brown et al (2020) note that spawner surveys have been very inconsistent and state that *"This may no longer be a distinct CU due to small population size, straying from Middle Shuswap and hatchery practices."* Therefore, the CU was not included in the retrospective test.
- *Lower Thompson_SP_1.2* (**Figure 66**): The integrated status assessment was *Red* with data up to 2012. All seven algorithms could assign status from 2009 to 2013, and all assign *Red* for those years, matching the integrated assessment in 2012. From 2014 to 2018, five algorithms could assign a status, with statuses mostly stable for each algorithm, but ranging from *Red* to *Green* across algorithms. All seven algorithms switched back to *Red* for 2019.
- *North Thompson_SP_1.3* (No Figure): Experts at the status workshop (DFO 2016) assigned *Red* status, noting "*very strong short-term decline and very low numbers of fish, combined with high uncertainty due to small number of data points."* Brown et al (2020) conclude that standard WSP metrics cannot be calculated for the available data due to poor counting conditions and inconsistent site coverage.

- *North Thompson_SU_1.3* (**Figure 67**): The integrated status assessment was *Red* with data up to 2012. All seven algorithms could assign status for all years starting in 2011 and assigned *Red* for all years with completed statuses, including 2012 to match the expert assessment.
- *Upper Adams River_SU_x.x* (No Figure): Experts at the status workshop (DFO 2016) designated this CU as data deficient, because available spawner estimates are based on redd counts, which are difficult to assess consistently, and the CU is not routinely surveyed. Brown et al (2020) further note that Chinook in the Upper Adams were extirpated by a dam, then re-stocked from various sources with different life histories after dam removal. While the population appears to be self-sustaining, but abundance estimates are low and data are sparse. Therefore, the CU was not included in the retrospective test.

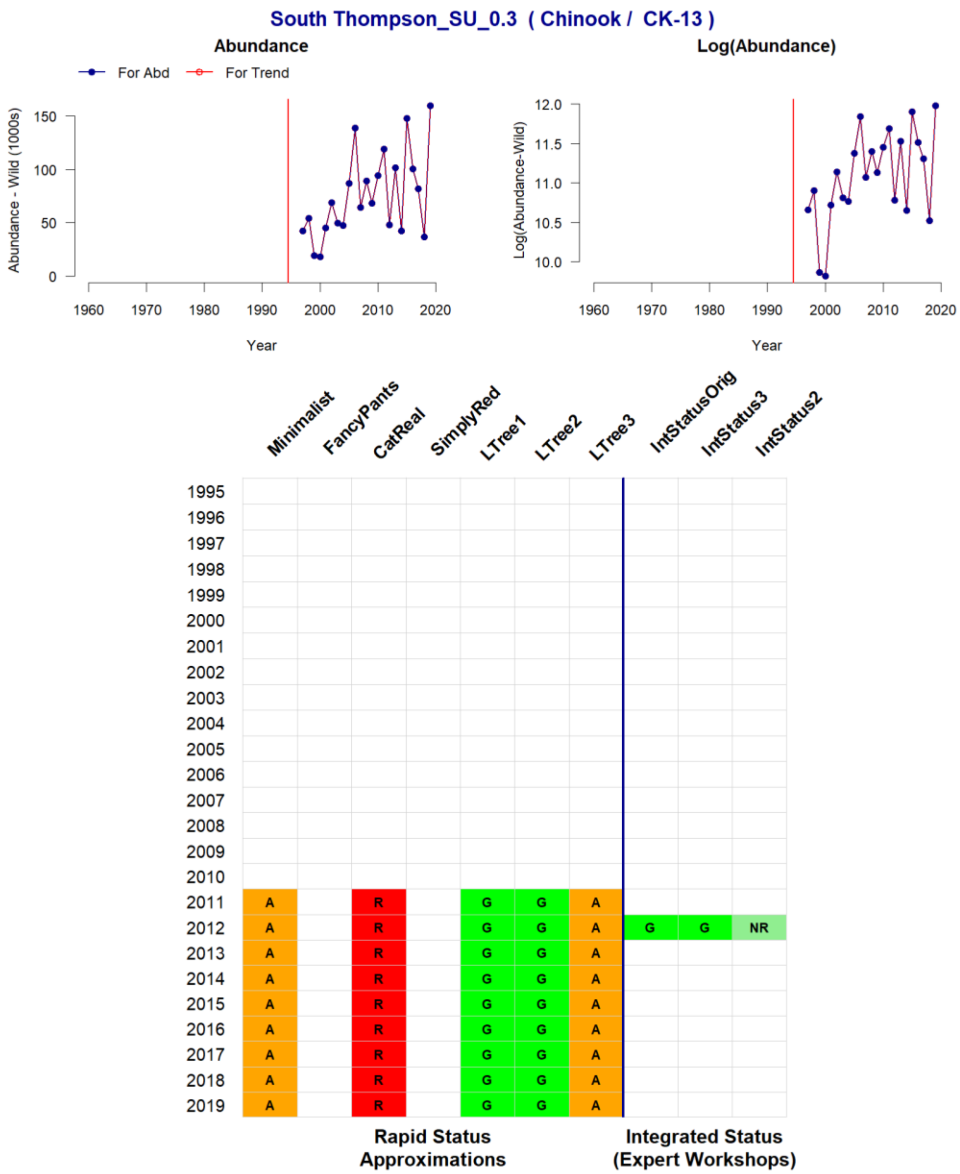


Figure 63. Retrospective test of rapid status approximations – Thompson Chinook – South Thompson_SU_0.3. Figure shows annual rapid statuses assigned by seven alternative candidate algorithms, as well as completed integrated status assessments. Note that integrated status assessments are mapped onto the year of data used, not the year the WSP status assessment was done. Statuses are denoted as R = *Red*, RA = *Red/Amber*, A = *Amber*, AG = *Amber/Green*, G = *Green*, NR = *NotRed*, DD = *Data Deficient*, and UD = *Undetermined*. For comparison, integrated statuses are shown as the original status assignment from the expert workshops (IntStatusOrig), as well as the conversions to the 3-status (*Red/Amber/Green*) and 2-status scales (*Red/NotRed*).

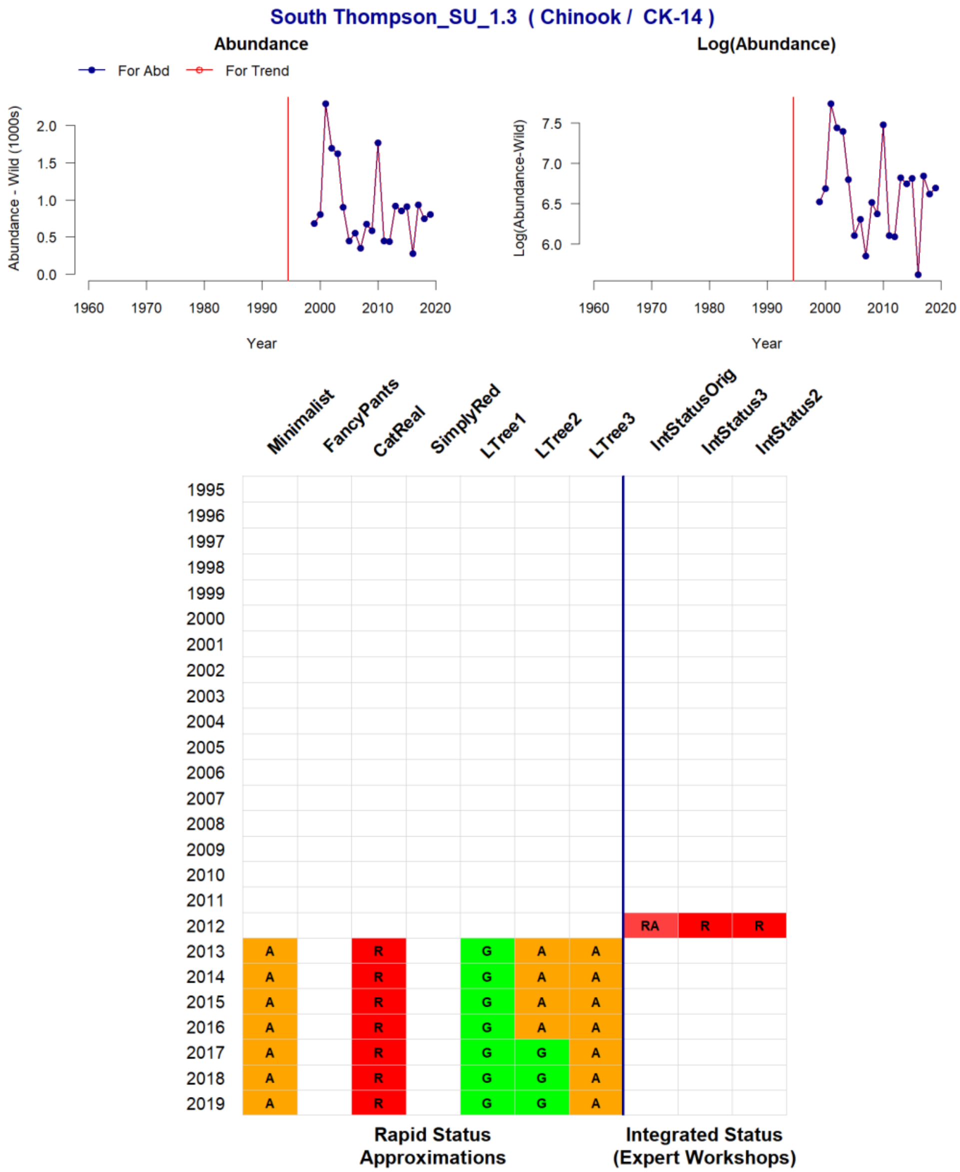


Figure 64. Retrospective test of rapid status approximations - Thompson Chinook – South Thompson_SU_1.3. Figure shows annual rapid statuses assigned by seven alternative candidate algorithms, as well as completed integrated status assessments. Note that integrated status assessments are mapped onto the year of data used, not the year the WSP status assessment was done. Statuses are denoted as R = *ed*, RA = *Red/Amber*, A = *Amber*, AG = *Amber/Green*, G = *Green*, NR = *NotRed*, DD = *Data Deficient*, and UD = *Undetermined*. For comparison, integrated statuses are shown as the original status assignment from the expert workshops (IntStatusOrig), as well as the conversions to the 3-status (*Red/Amber/Green*) and 2-status scales (*Red/NotRed*).

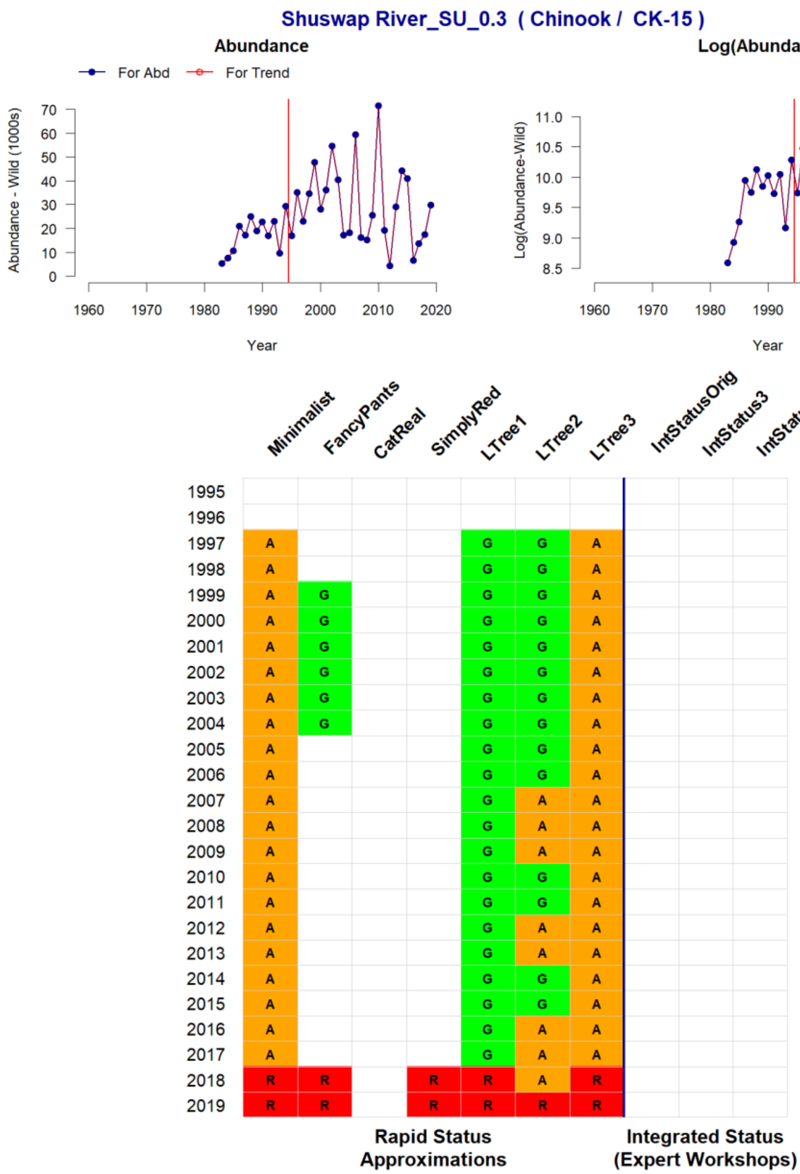


Figure 65. Retrospective test of rapid status approximations - Thompson Chinook – Shuswap River_SU_0.3. Figure shows annual rapid statuses assigned by seven alternative candidate algorithms, as well as completed integrated status assessments. Note that integrated status assessments are mapped onto the year of data used, not the year the WSP status assessment was done. Statuses are denoted as R = *Red*, RA = *Red/Amber*, A = *Amber*, AG = *Amber/Green*, G = *Green*, NR = *NotRed*, DD = *Data Deficient*, and UD = *Undetermined*. For comparison, integrated statuses are shown as the original status assignment from the expert workshops (IntStatusOrig), as well as the conversions to the 3-status (*Red/Amber/Green*) and 2-status scales (*Red/NotRed*).

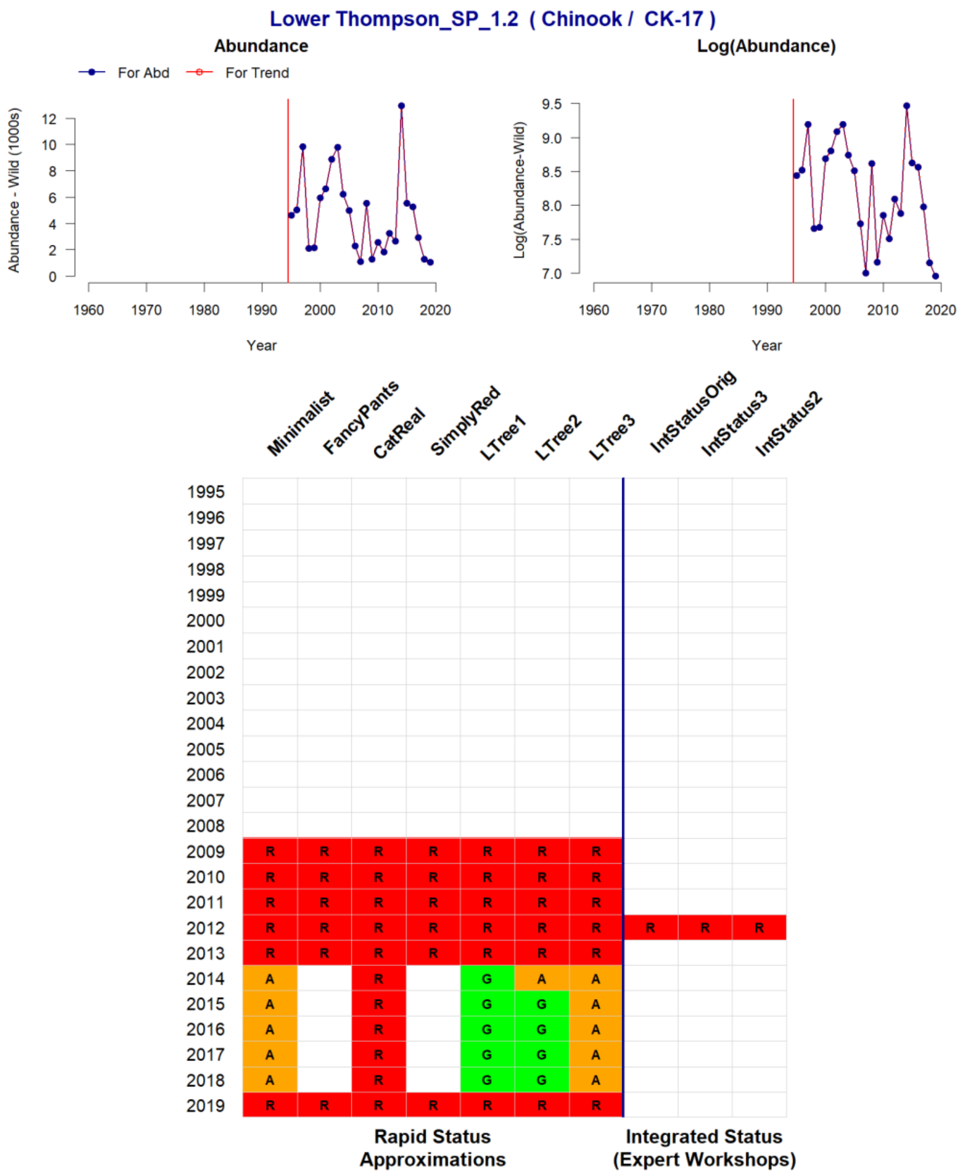
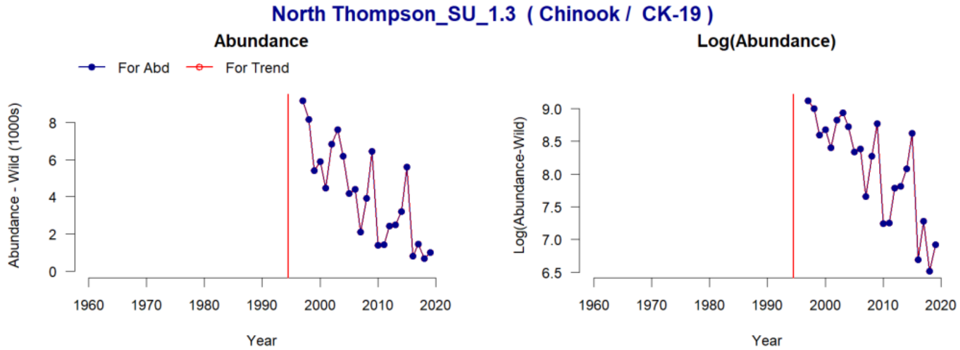


Figure 66. Retrospective test of rapid status approximations - Thompson Chinook – Lower Thompson_SP_1.2. Figure shows annual rapid statuses assigned by seven alternative candidate algorithms, as well as completed integrated status assessments. Note that integrated status assessments are mapped onto the year of data used, not the year the WSP status assessment was done. Statuses are denoted as R = *Red*, RA = *Red/Amber*, A = *Amber*, AG = *Amber/Green*, G = *Green*, NR = *NotRed*, DD = *Data Deficient*, and UD = *Undetermined*. For comparison, integrated statuses are shown as the original status assignment from the expert workshops (IntStatusOrig), as well as the conversions to the 3-status (*Red/Amber/Green*) and 2-status scales (*Red/NotRed*).



Year	Minimalist	FancyPants	CatReal	SimplyRed	LTree1	LTree2	LTree3	IntStatusOrig	IntStatus3	IntStatus2
1995										
1996										
1997										
1998										
1999										
2000										
2001										
2002										
2003										
2004										
2005										
2006										
2007										
2008										
2009										
2010										
2011	R	R	R	R	R	R	R			
2012	R	R	R	R	R	R	R	R	R	R
2013	R	R	R	R	R	R	R			
2014	R	R	R	R	R	R	R			
2015	R	R	R	R	R	R	R			
2016	R	R	R	R	R	R	R			
2017	R	R	R	R	R	R	R			
2018	R	R	R	R	R	R	R			
2019	R	R	R	R	R	R	R			

Figure 67. Retrospective test of rapid status approximations - Thompson Chinook – North Thompson_SU_1.3. Figure shows annual rapid statuses assigned by seven alternative candidate algorithms, as well as completed integrated status assessments. Note that integrated status assessments are mapped onto the year of data used, not the year the WSP status assessment was done. Statuses are denoted as R = *Red*, RA = *Red/Amber*, A = *Amber*, AG = *Amber/Green*, G = *Green*, NR = *NotRed*, DD = *Data Deficient*, and UD = *Undetermined*. For comparison, integrated statuses are shown as the original status assignment from the expert workshops (IntStatusOrig), as well as the conversions to the 3-status (*Red/Amber/Green*) and 2-status scales (*Red/NotRed*).

H.11 SOUTHERN BC CHINOOK – INNER SOUTH COAST

Full integrated status assessments of Thompson Chinook were completed in 2012 (DFO 2016), using spawner data up to 2012. The integrated assessment and retrospective test cover 8 CUs in the Inner South Coast Chinook management unit:

- *Boundary Bay_FA_0.3* (No Figure):
- *Southern Mainland-Georgia Strait_FA_0.x* (No Figure): Experts at the status workshop (DFO 2016) designated this CU as data deficient, because "*No recent, high quality escapement records for wild sites.*" Brown et al (2020) summarize the data issues. Therefore, the CU was not included in the retrospective test.
- *Southern Mainland-Southern Fjords_FA_0.x* (No Figure): Good quality data available for one highly enhanced site, but none for wild sites. Therefore, the CU was not included in the retrospective test. Brown et al (2020) summarize the data issues.
- *East Vancouver Island-Goldstream_FA_0.x* (No Figure): No data for wild sites. Therefore, the CU was not included in the retrospective test. Brown et al (2020) summarize the data issues.
- *East Vancouver Island-Cowichan & Koksilah_FA_0.x* (No Figure): No data for wild sites. Therefore, the CU was not included in the retrospective test. Brown et al (2020) summarize the data issues.
- *East Vancouver Island-Nanaimo_SP_1.x* (No Figure): No enhancement, but insufficient data for status assessment. Therefore, the CU was not included in the retrospective test. Brown et al (2020) summarize the data issues.
- *East Vancouver Island-Nanaimo & Chemainus_FA_0.x* (No Figure): No data for wild sites. Therefore, the CU was not included in the retrospective test. Brown et al (2020) summarize the data issues.
- *East Vancouver Island-Qualicum & Puntledge_FA_0.x* (No Figure): No data for wild sites. Therefore, the CU was not included in the retrospective test. Brown et al (2020) summarize the data issues.
- *East Vancouver Island-North_FA_0.x* (Figure 68): The integrated status assessment was *Red* with data up to 2012. Five of the seven algorithms could assign status starting in 2013, the other two algorithms could only assign status for 2015 and 2016. For those two years, all 7 algorithms assigned *Red*, but for other years the algorithm results range from *Red* to *Green*. *Learning Tree 3* assigns *Amber* for years other than 2015 and 2016.

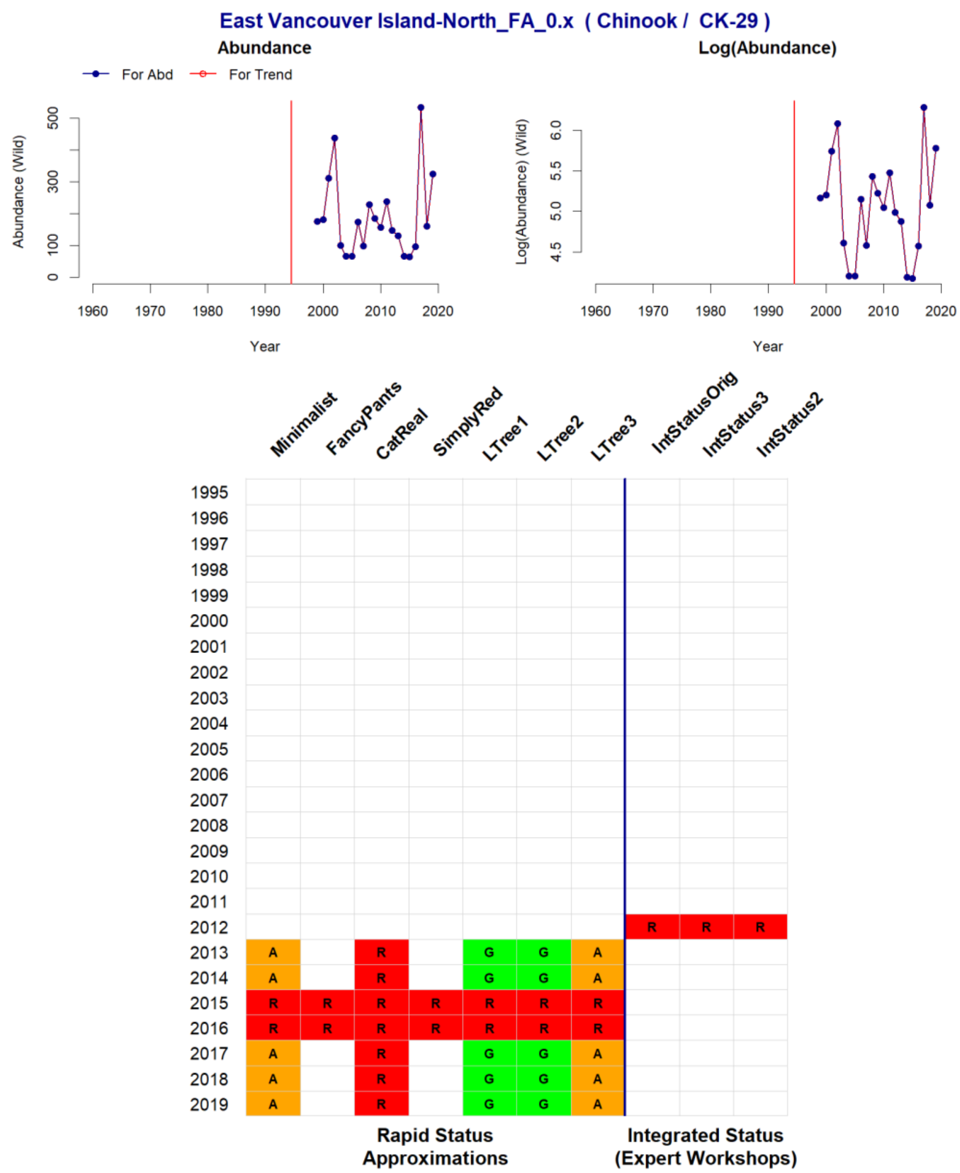


Figure 68. Retrospective test of rapid status approximations – Inner South Coast Chinook – East Vancouver Island-North_FA_0.x. Figure shows annual rapid statuses assigned by seven alternative candidate algorithms, as well as completed integrated status assessments. Note that integrated status assessments are mapped onto the year of data used, not the year the WSP status assessment was done. Statuses are denoted as R = *Red*, RA = *Red/Amber*, A = *Amber*, AG = *Amber/Green*, G = *Green*, NR = *NotRed*, DD = *Data Deficient*, and UD = *Undetermined*. For comparison, integrated statuses are shown as the original status assignment from the expert workshops (IntStatusOrig), as well as the conversions to the 3-status (*Red/Amber/Green*) and 2-status scales (*Red/NotRed*).

H.12 SOUTHERN BC CHINOOK – WEST COAST VANCOUVER ISLAND

Full integrated status assessments of WCVI Chinook were completed in 2012 (DFO 2016) , using spawner data up to 2012. The integrated assessment and retrospective test covers 3 CUs in the WCVI Chinook management unit:

- *WCVI-South-Fall 0.x* (Figure 69): The integrated status assessment was *Red* with data up to 2012. Four of the seven algorithms could assign statuses for every year since 1995. The other three algorithms could assign a status for every year since 2009. For five algorithms, all the assigned statuses are *Red*. Six of the seven algorithms assign *Red* for 2012, matching the integrated status assign in the expert workshop. Notes from the workshop in 2012 state: "Most groups designated this CU as *Red*, but due mostly to pressures (straying from large-scale hatchery releases, including seapens, and high exploitation rates (roughly 60%) rather than to abundance or observed trends. Data from 2 small populations among 21 possible wild sites is not considered to be representative. Participants recommended completion of further work to determine whether these populations still exist as a CU under WSP definition". Note that subsequent data revisions increased the number of indicator systems included in the time series to four.
- *WCVI-Nootka & Kyuquot-Fall 0.x* (Figure 70): The integrated status assessment was *Red* with data up to 2012. Four of the seven algorithms could assign statuses for every year since 1998. The other three algorithms could assign a status for every year since 2009. Six of the algorithms show status dropping to *Red* for several years around 2010, then starting to improve around 2015. Notes from the expert workshop in 2012 state: "Most groups designated this CU as *Red*, but this was the result of considerations other than the 3 WSP metrics. Rather, participants highlighted the following concerns: only a small portion of total abundance in wild sites and impacts of straying are likely, very small index of abundance of wild sites."
- *WCVI-North-Fall 0.x* (Figure 71): The CU was not assessed in the expert workshop, but subsequent revisions to site classifications generated a CU-level time series based on two indicator systems (Marble, Cayeghle). Four of the seven algorithms assign a consistent status since 1999, but statuses differed between algorithms (*Categorical Realist* and *Learning Tree 3* assign Amber to all years, *Learning Trees* 1 and 2 assign *Green* to all years). The other three algorithms assign *Amber/NotRed* starting in 2010, then shift to *Red* from 2013-2016, and back to *Amber/NotRed* in 2017.

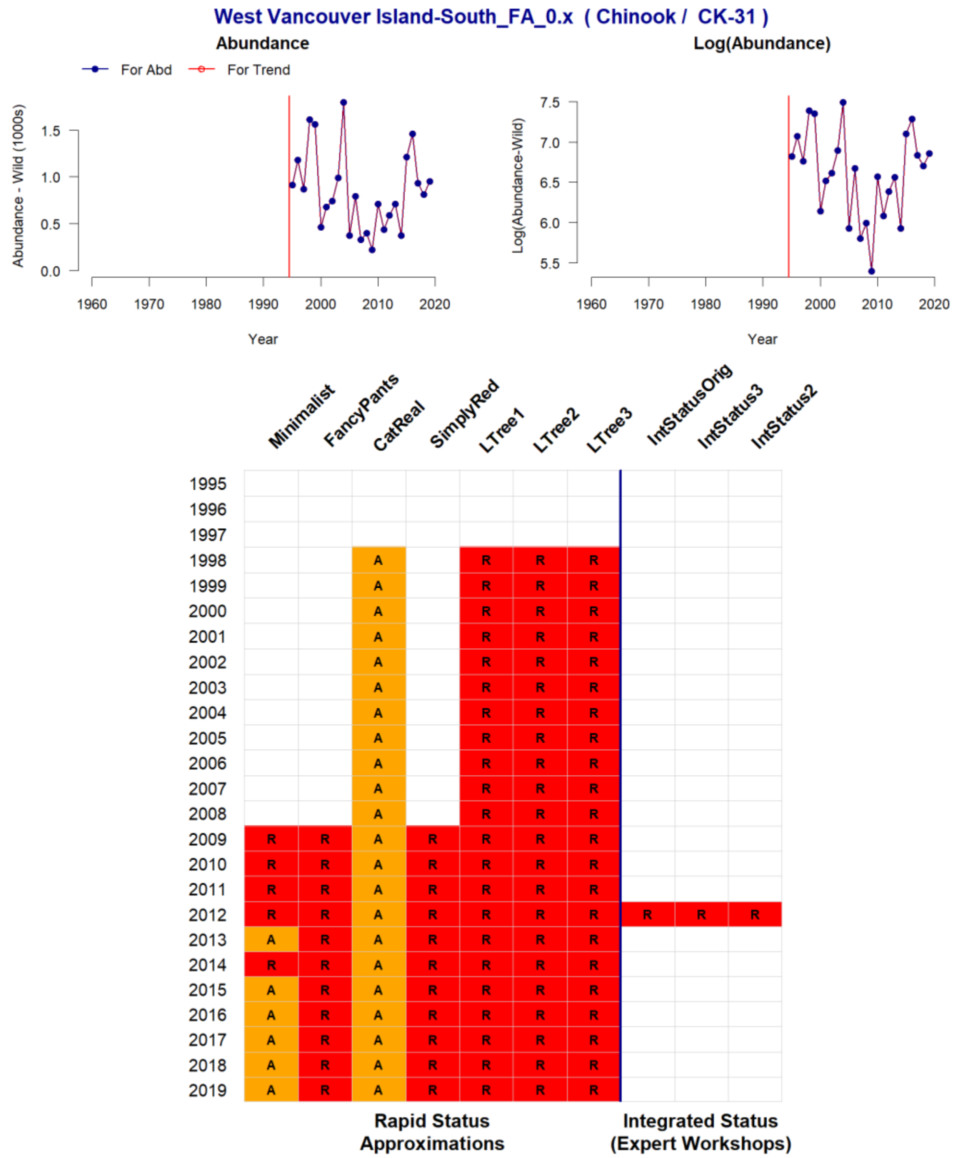
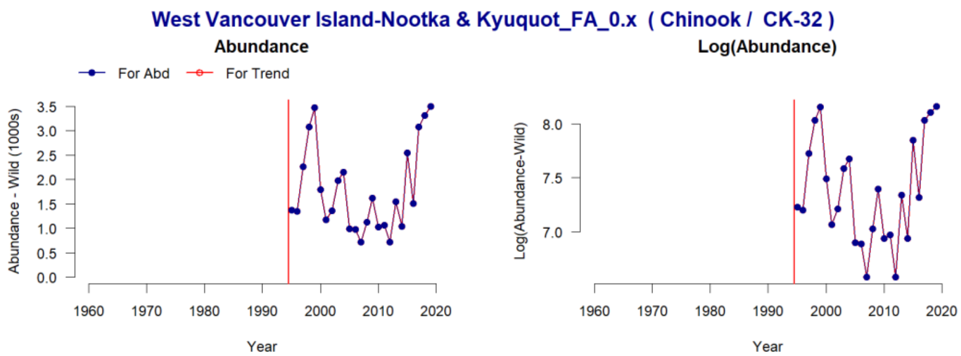


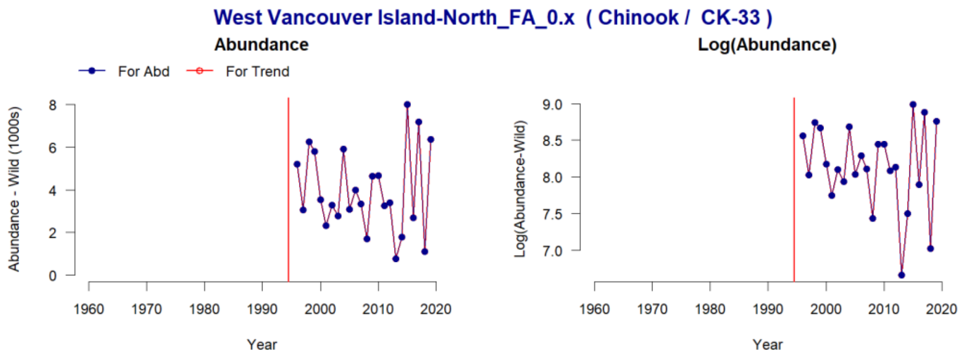
Figure 69. Retrospective test of rapid status approximations – SBC Chinook – WCVI-South-Fall 0.x. Figure shows annual rapid statuses assigned by seven alternative candidate algorithms, as well as completed integrated status assessments. Note that integrated status assessments are mapped onto the year of data used, not the year the WSP status assessment was done. Statuses are denoted as R = *Red*, RA = *Red/Amber*, A = *Amber*, AG = *Amber/Green*, G = *Green*, NR = *NotRed*, DD = *Data Deficient*, and UD = *Undetermined*. For comparison, integrated statuses are shown as the original status assignment from the expert workshops (IntStatusOrig), as well as the conversions to the 3-status (*Red/Amber/Green*) and 2-status scales (*Red/NotRed*).



	Minimalist	FancyPants	CatReal	SimplyRed	LTree1	LTree2	LTree3	IntStatusOrig	IntStatus3	IntStatus2
1995										
1996										
1997										
1998			A		A	A	A			
1999			A		A	G	A			
2000			A		G	G	A			
2001			A		A	A	A			
2002			A		A	A	A			
2003			A		A	A	A			
2004			A		A	A	A			
2005			A		A	A	A			
2006			A		A	A	R			
2007			A		R	R	R			
2008			A		R	R	R			
2009	R	R	A	R	R	R	R			
2010	R	R	A	R	R	R	R			
2011	A	A	A	R	R	R	R			
2012	R	R	A	R	R	R	R	R	R	R
2013	R	R	A	R	R	R	R			
2014	R	R	A	R	R	R	R			
2015	A	A	A	NR	A	A	R			
2016	A	AG	A	NR	A	A	A			
2017	A	AG	A	NR	A	A	A			
2018	A	AG	A	NR	G	G	A			
2019	A	AG	A	NR	G	G	A			

Rapid Status Approximations — Integrated Status (Expert Workshops)

Figure 70. Retrospective test of rapid status approximations – SBC Chinook – WCVI-Nootka & Kyuquot-Fall 0.x. Figure shows annual rapid statuses assigned by seven alternative candidate algorithms, as well as completed integrated status assessments. Note that integrated status assessments are mapped onto the year of data used, not the year the WSP status assessment was done. Statuses are denoted as R = *Red*, RA = *Red/Amber*, A = *Amber*, AG = *Amber/Green*, G = *Green*, NR = *NotRed*, DD = *Data Deficient*, and UD = *Undetermined*. For comparison, integrated statuses are shown as the original status assignment from the expert workshops (IntStatusOrig), as well as the conversions to the 3-status (*Red/Amber/Green*) and 2-status scales (*Red/NotRed*).



	Minimalist	FancyPants	CatReal	SimplyRed	LTree1	LTree2	LTree3	IntStatusOrig	IntStatus3	IntStatus2
1995										
1996										
1997										
1998										
1999			A		G	G	A			
2000			A		G	G	A			
2001			A		G	G	A			
2002			A		G	G	A			
2003			A		G	G	A			
2004			A		G	G	A			
2005			A		G	G	A			
2006			A		G	G	A			
2007			A		G	G	A			
2008			A		G	G	A			
2009			A		G	G	A			
2010	A	A	A	NR	G	G	A			
2011	A	A	A	NR	G	G	A			
2012	A	A	A	NR	G	G	A			
2013	R	R	A	R	G	G	A			
2014	R	R	A	R	G	G	A			
2015	R	R	A	R	G	G	A			
2016	R	R	A	R	G	G	A			
2017	A	A	A	NR	G	G	A			
2018	A	A	A	NR	G	G	A			
2019	A	A	A	NR	G	G	A			

Rapid Status Approximations | Integrated Status (Expert Workshops)

Figure 71. Retrospective test of rapid status approximations – SBC Chinook – WCVI-Nootka & Kyuquot-Fall 0.x. Figure shows annual rapid statuses assigned by seven alternative candidate algorithms, as well as completed integrated status assessments. Note that integrated status assessments are mapped onto the year of data used, not the year the WSP status assessment was done. Statuses are denoted as R = *Red*, RA = *Red/Amber*, A = *Amber*, AG = *Amber/Green*, G = *Green*, NR = *NotRed*, DD = *Data Deficient*, and UD = *Undetermined*. For comparison, integrated statuses are shown as the original status assignment from the expert workshops (IntStatusOrig), as well as the conversions to the 3-status (*Red/Amber/Green*) and 2-status scales (*Red/NotRed*).

H.12 SOUTHERN BC CHINOOK – OKANAGAN

Full integrated status assessments of WCVI Chinook were completed in 2012 (DFO 2016) , using spawner data up to 2012. The integrated assessment and retrospective test covers 1 CUs in the Okanagan Chinook management unit:

- *Okanagan-1.x (Figure 72):* The integrated status assessment was *Red* with data up to 2012. The data was completely revised since then, but *Learning Trees 1-3* assigned *Red* for all years starting in 2009, matching the expert assessment for 2012. *Minimalist* could assign status for 2 years.

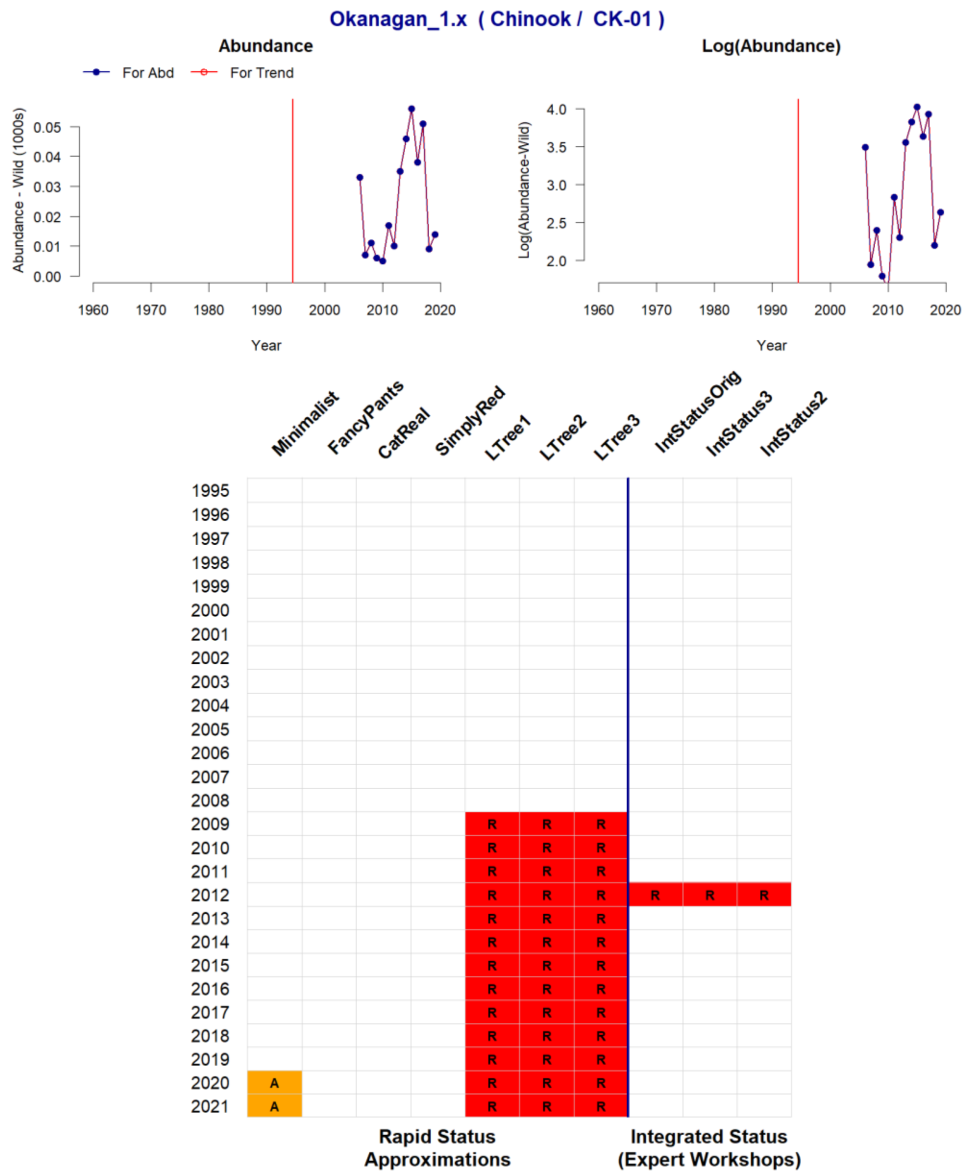


Figure 72. Retrospective test of rapid status approximations – Okanagan Chinook – Okanagan_1.x. Figure shows annual rapid statuses assigned by seven alternative candidate algorithms, as well as completed integrated status assessments. Note that integrated status assessments are mapped onto the year of data used, not the year the WSP status assessment was done. Statuses are denoted as R = *Red*, RA = *Red/Amber*, A = *Amber*, AG = *Amber/Green*, G = *Green*, NR = *NotRed*, DD = *Data Deficient*, and UD = *Undetermined*. For comparison, integrated statuses are shown as the original status assignment from the expert workshops (IntStatusOrig), as well as the conversions to the 3-status (*Red/Amber/Green*) and 2-status scales (*Red/NotRed*).