Improving pre-season planning and in-season estimates of Fraser River sockeye stocks through stock- and cycle line-specific estimates

2020 Annual Report to the Southern Fund Committee

March 2020

Angela Phung, Catherine Michielsens, Merran Hague and Mark McMillan.

Pacific Salmon Commission

Background

Historical data generated from reconstructed estimates of daily marine abundances of Fraser River sockeye salmon form the basis for critical components of pre-season planning and in-season stock assessments. Pre-season, historical time-series of timing and diversion are used to parameterise forecast models and fisheries simulations (Cave and Gazey 1994). In-season, reconstructions are used to update estimates of run abundance, timing, diversion rate and test fishery catchability (expansion lines, Michielsens and Cave 2019). For fisheries planning and assessment purposes, PSC Secretariat biologists assess these estimates at a refined stock resolution, often subset by cycle-line. Over time, the stock-resolution, file structure, and assumptions applied to in-season run reconstructions has evolved, leading to inconsistencies across the time series. In addition, post-season adjustments, including updated stock ID and corrected/updated catches, can lead to a disparity between the in-season run reconstruction files and post-season estimates.

In 2018, PSC Secretariat staff submitted a two-year project proposal to update and standardise historical run reconstruction files using recently updated post-season catch and Mission passage information and currently adopted data standards (Patterson et al. 2017 - S16-I03 & S15-I11: Improvements to predicting en-route loss estimates for Fraser sockeye salmon 2015-16). We also proposed to build a new data warehouse to store post-season estimates of daily run reconstructions, from which time series of fish migration behaviour and fishery harvest parameters can easily be updated and queried at various stock resolutions and year-aggregations (e.g. by cycle line or dominant year). These improvements will bring the quality of the run reconstruction estimates in line with the recent improvements to the post-season catch, passage and escapement records, and help the Secretariat provide the Fraser River Panel with key inputs to both the pre-season and in-season management processes.

The overall project has been broken down into 3 different steps: 1) standardisation of post-season run reconstruction model files for 1996 – 2017, 2) development and population of a run reconstruction database and 3) updating time series of migration timing, spread, diversion and Gulf delay. During the first year of the project, progress has been made on the first two steps of the project.

Progress report

As identified in the project schedule (see attached), the following work items were assigned to Year 1 of this project: 1) development of an Excel template and application to other years, 2) updating of links to the RuFEs data, 3) standardisation of model assumption and 4) the design and construction of the database architecture (to be continued in Year 2). Items 1-3 fall under the first steps within the database project, i.e. standardisation of the post-season run reconstruction files for 1996-2017, while items 4 falls under step 2, i.e. development and population of a run reconstruction database.

Using 2018 as a case study, the development of a full suite of Excel templates (Work Item 1) is nearly completed but more testing and fine tuning of the templates is required. In the last interim report several advancements had been identified and these advancements have now been incorporated: inclusion of higher-resolution catch and racial data produced as part of the SEF project S16-I03 (Patterson et al. 2017, Work Item 2), increased user flexibility to accommodate different model assumption, and a framework to explicitly reconstruct fish abundance by migration route, i.e. in Johnstone Strait versus Juan de Fuca Strait.

Incorporating high-resolution catch and racial data into the marine reconstruction files has led to larger and slower files; therefore, different options to optimize and speed up these files were explored. In part, the creation of these new Excel templates has been facilitated by VBA code which may take longer to develop in the beginning but will greatly speed up the process of applying the Excel templates to other historical years of data. Initial work has begun on applying a subset of the Excel templates to other historical years (Work Item 1) but has required some trouble shooting as different years of data contain different file structures which affects how files link to each other. Although the Excel templates are nearly complete, the application of these templates to other historical years will continue into Year 2 of the project.

Time was dedicated to standardising timing offsets and spread assumptions across different files (Work Item 3). A literature review was conducted to obtain standard swim speeds by approach route, and a spatial information database (QGIS) of migration distances was created to standardize sockeye travel assumptions as they swim through fishery management areas. Differences in timing offsets between different files has been documented and the next step will be to implement offset

changes across different files and working groups once changes have been approved by managers and biologists.

Despite ongoing work on developing the Excel templates, there are already preliminary high-resolution daily abundance outputs by migration route and location available for 2018. A process was developed to seamlessly export Excel data into a database-friendly output which can be used to build and populate the DARMA (Daily Reconstructed Marine Abundance) database (Work Item 4). Work can soon begin on DARMA's architecture, implementation, and system development by leveraging existing RUFES database architecture and relational entities. The new excel templates will need to be applied to historical data before all the data are available to be imported into DARMA. Therefore, the timeline for importing all historical data in DARMA will likely extend into the Winter of 2020.

Next steps

The following work items will further support the development and population of a run reconstruction database and will be the focus in Year 2 of this project (i.e. anticipated completion is on or before April 1, 2021): 1) creation of a cloud-based relational database, 2) development of a user interface, 3) development of a data migration process and 4) establishing queries for historical time series.

Acknowledgements

We would like to thank the Pacific Salmon Commission, Southern Restoration and Endowment Fund Committee for their support and funding of this project.

References

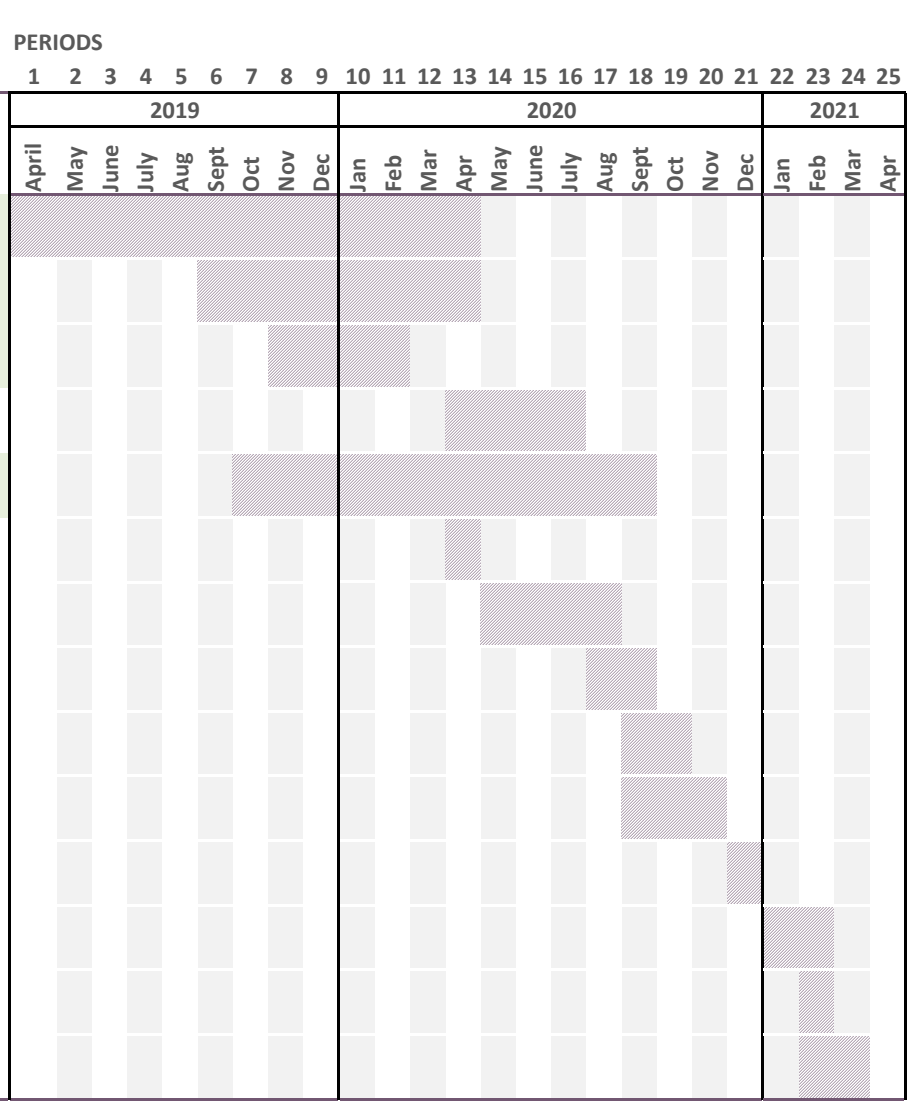
Cave, J.D. and W.J. Gazey. 1994. A Preseason Simulation Model for Fisheries on Fraser River Sockeye Salmon (*Oncorhynchus nerka*). CJFAS 51: 1535-1549.

Michielsens, C.G.J. and J.D. Cave. 2019. In-season assessment and management of salmon stocks using a Bayesian time-density model. CJFAS 76: 1073-1085.

Patterson, D.A., Guthrie, I., Lapointe, M.L., Huang, A-M., Braun, D., and S.G. Hinch. 2017. S16-I03 & S15-I11: Improvements to predicting en-route loss estimates for Fraser sockeye salmon 2015-16.

Michielsens 2020 - Reconstruction Database Project Schedule

ACTIVITY	PLAN START	PLAN DURATION
Develop Excel template[1]	1	13
Update links to RuFEs data[1]	6	8
Standardise model assumptions[1]	8	4
Validate and document differences in historical vs. updated data	13	4
Design/build database architecture[1]	7	12
Write interim report	13	1
Design/build database input interface	14	4
Upload reconstruction historical data and run QA/QC	17	2
Migrate data from Access to Azure	18	2
Design/build database export interface & queries	18	3
Query updated timeseries	21	1
Validate and document differences in historical vs. updated time series	22	2
Present results to Fraser River Panel	23	1
Write Final Report	23	2



[1]Year 1 work items; note that included in work item "Develop Excel template" is the task of applying the template to other years.