Improving the Kuskokwim In-Season Salmon Harvest Estimation Workflow

Project Summary Report

Benjamin A. Staton
Quantitative Ecological Services LLC, Portland, OR

Project Summary

Methods to produce in-season estimates of subsistence salmon harvest and effort for the lower Kuskokwim River have existed since 2016 (Staton and Coggins 2016). However, generating the estimates and reporting documents from the raw data has, until the 2021 season, required specialized knowledge of writing and editing computer code. The previous workflow required much experience with the code to implement it successfully, and in some cases the analyst was required to make subjective decisions about which data should be considered unreliable. Further, installing all necessary software was cumbersome and code and data files needed to be organized in highly specific (and often non-intuitive) ways. These undesirable features of the workflow existed because the same student/employee served as developer and analyst – producing the estimates in-season did not leave time for such optimization. After 5 years of implementation, however, it has become clear that some features of the workflow can be automated and better-organized, and that an overall more intuitive implementation is possible.

The primary objective of this project was to enable someone to produce estimates and reporting documents using the same analytical methods as in the past, but without needing to write or edit code. Other objectives included (*i*) automation of previously subjective choices about data quality using agreed-upon rules, (*ii*) simplification of the software installation process, and (*iii*) improvements to the software documentation resources. These objectives were attained through the development of a new software package called 'KuskoHarvEst' (Staton 2021) for program R (R Core Team 2020), which is the code environment where all calculations take place. The 'KuskoHarvEst' package performs all of the fundamental calculations that previously existed, but vastly streamlines their implementation for the user by providing a set of graphical user interfaces (i.e., menu-driven, point-and-click, text-entry boxes, etc.) that remove the need to edit any code by hand. These features increase the ease and speed of producing estimates, as well as improves their reproducibility among multiple analysts faced with the same data set.

'KuskoHarvEst' was used successfully in 2021 by staff from the Kuskokwim River Inter-Tribal Fish Commission and the Orutsararmiut Native Council to produce harvest and effort estimates and distribute summaries in the form of report documents with minimal oversight from the software developer. Feedback gathered from the 2021 season was used post-season to fix minor issues and add features that have lead to a finalized version (*v*1.1.0) that can be used in future years, assuming the monitoring program and fishery conditions remain relatively consistent.

Project Deliverables

This project produced several tangible deliverables (e.g., software, documents, presentation slides) as well as some that are less tangible (e.g., training, oversight, and data checking).

1. The 'KuskoHarvEst' R package, available indefinitely at https://www.github.com/bstaton/KuskoHarvEst and also delivered in zip file format. The package contains all source code to complete calculations and produce reports via graphical user interfaces. Full instructions to install all needed resources are available at the link.

2. Comprehensive documentation to train new users (and refresh experienced users) in how to successfully use the software. This content is technically built into 'KuskoHarvEst', but given the extent of it, it is listed here as a separate deliverable.

   General instructions for the different parts of the workflow.

   Description of raw data file formatting required by the software.

   Instructions for the meta-data tool where users enter features of the fishing day like allowed time, area, and gear information.

   Instructions for the interview/flight data preparation tool where users load the raw data files and automatically screen the data for potential problems.

   Instructions for the report builder tool where users select which content to include in the report and click one button to to perform all calculations and generate a standardized PDF report document that contains the results and data summaries based on the input data.

   Instructions to generate content for the season-wide final report including figures and tables that follow the formatting of this content from previous years' reports (Staton and Coggins 2016, 2017; Staton 2018; Decossas 2019, 2020).

   A description of all quality assurance checks performed automatically by the software that decide which interview records are suitable for use in each of the tasks in the estimation workflow.

   A description of the coding framework used by the software, including (*i*) a demonstration of how to produce estimates without the interactive tools, (*ii*) a description of the role of each custom R function, and (*iii*) an overview of how the output of calculations are automatically inserted into PDF report documents.

3. Slides from a presentation describing the statistical foundations and assumptions of the harvest and effort estimators (presented on May 18, 2021).

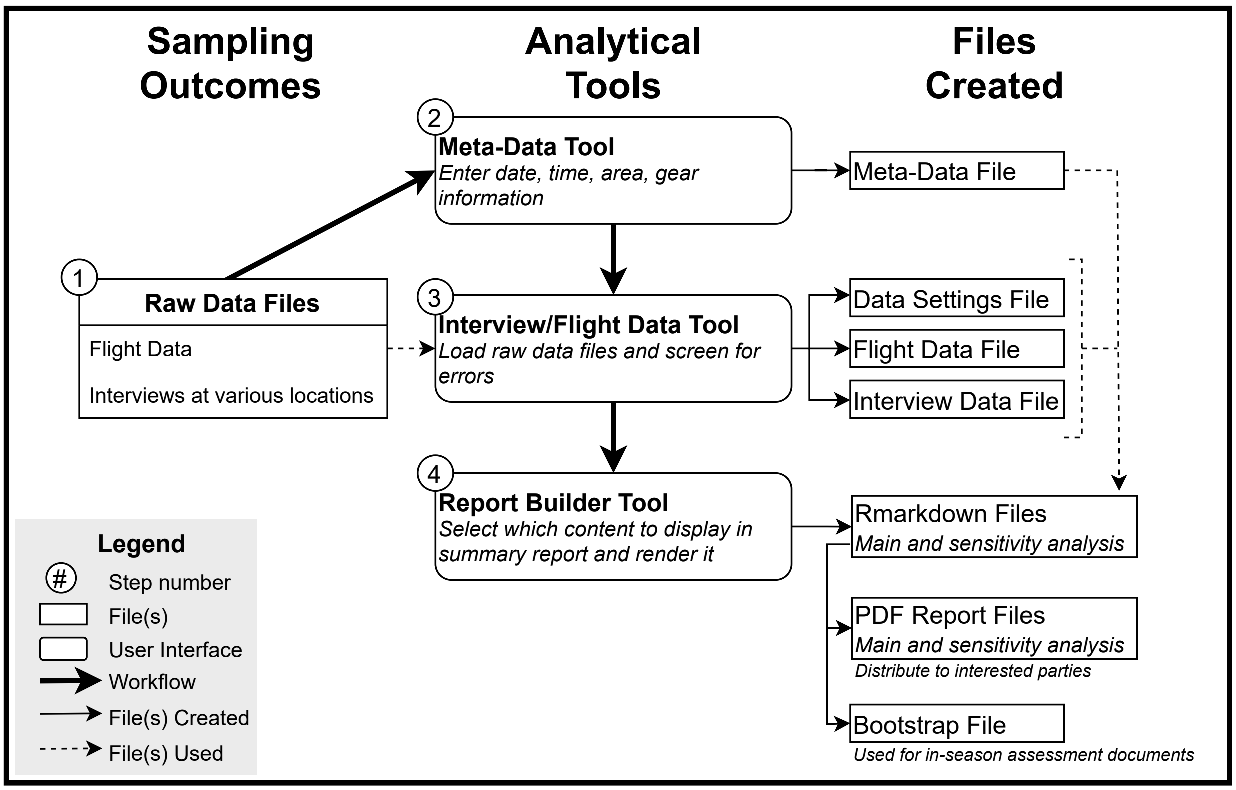
4. Training and in-season oversight for the analysts using the software to produce estimates, which helped to quickly identify data entry issues and give the analysts confidence that they used the software properly. The necessity of oversight generally declined as the season progressed.

5. Assistance with the final report (Russell et al. 2021) that documents the data availability for the 2021 season, raw data summaries, and summaries of compiled harvest and effort estimates.

Workflow Description

This section describes the steps carried out by the analyst to produce harvest and effort estimates using the 'KuskoHarvEst' package – Figure 1 shows a visual schematic. The 'KuskoHarvEst' workflow begins following the conclusion of a subsistence harvest opportunity in which both on-the-ground interviews and aerial surveys were conducted.

FIGURE 1. Layout of the 'KuskoHarvEst' workflow showing the steps involved in generating harvest and effort estimates for one day of fishing as well as the files created in the process.



1. The analyst must **compile all data** that were gathered during the opportunity – great care must be taken during this step since, by design, errors in raw data entry or formatting are the source of any issues the analyst may encounter later in the workflow (complete details on data file formatting can be found here). After the data have been compiled, the analyst creates a new folder on their computer (aided by 'KuskoHarvEst') to serve as the location where the software will automatically look for and export necessary files – the analyst must place the data files in this location manually.

2. The analyst next must access the **"Meta-Data Tool"**, which is a simple interactive interface (see here for visuals and instructions) where users enter information relevant to the fishing day such as the dates, times, and area coverage of the estimates, web links to the announcement of the fishing opportunity, contact information for the

analyst(s), and which types of fishing gear were allowed (i.e., set net only versus set and drift nets). Because of the interface design, it is impossible for users to incorrectly format their selections here, which is critical since many of the calculations in later steps depend on the information entered in this step. As such, this step creates a new file that stores the entries made by the user ("Meta-Data File"; Figure 1).

3. The analyst next must access the **"Interview/Flight Data Tool"**, which is a more complex, yet still intuitive, interactive interface (see here for visuals and instructions). Here, the user selects which files to include, sets input values for data screening rules (pre-populated with default values that are recommended to never change), loads the data which are automatically screened for issues, and views the outcomes of these data preparation steps: a compiled data set of all interviews that were conducted. Please see here for a complete description of all automated data checks made by the software. As part of this step, the user also instructs the software to prepare the flight data into a standardized format. If there are any data formatting errors, the program would crash at this point and the analyst would need to exit the interface and correct them before proceeding.

4. The next and final step for the analyst is to access the **"Report Builder Tool"**, where they select which content to include in the output report (via check boxes), build the "Rmarkdown file" (with the click of a button; more details below), and render the main PDF document (with the click of another button; see here for visuals and instructions). This PDF is the file that is distributed to other biologists, managers, and stakeholders that displays data summaries and estimates – see Figures A1 - A4 for examples of the pages of this document. Users also perform these tasks to produce the sensitivity analyses PDF document through this same tool, which can give useful insights on the reliability of the estimates (see Figures B1 – B3 for examples of the pages of this document). Also created in this step is the "Bootstrap File", which contains the harvest estimates and their uncertainty and is useful in calculating uncertainty in cumulative harvest estimates.

What is Rmarkdown?

Given its central role in enabling the automated generation of report documents, some additional details on Rmarkdown are warranted. Rmarkdown is a framework that enables generating output documents that report analytical results based on inputs where the analysis is embedded within the report itself. That is, the code that creates the document also contains the analysis, so that if the inputs change, the in-text numbers, tables, and figures in the output document change accordingly and automatically. This prevents the user from needing to edit the document by hand, they need only ever edit the inputs (i.e., data files). Example output pages from the June 19, 2021 estimate are shown in Figures A1 – A4; note how many numbers there are on each page, and envision how long it would take the analyst to update all of this content by hand if the analysis was not integrated within the document. Rmarkdown is itself an R package ('rmarkdown', Xie et al. 2018) supported and extended by many other R packages, e.g., 'knitr' (Xie 2020) and 'kableExtra' (Zhu 2021).

Specific Issues and Solutions

Although a workflow using Rmarkdown was previously available, it suffered from some usability issues. This section highlights the four primary undesirable features of the old framework (developed and used by B. Staton in 2016 – 2018 and used in 2019 and 2020 by G. Decossas) and how they were addressed by 'KuskoHarvEst'. With the exception of several standardized/automated decisions that were previously subjective/manual (see item 3, below), the core statistical framework described in Staton (2018) and used between 2016 – 2020 is still contained in 'KuskoHarvEst'.

Nearly all of the solutions described below were enabled by bundling the estimation and reporting software into an R package.

(1) Copy/Paste/Edit Workflow

As originally developed, the estimation/reporting workflow required copying/pasting the code from a previous estimate and editing aspects of it to be suitable for the current estimate. This required knowledge of the code and which parts needed to be edited, but even for experienced users this could be time-consuming and error-prone depending on the needed changes. In some simple cases, for example, the only things that required editing were dates and file names. In other cases, the code needed to be altered by hand based on (*i*) how many flights were conducted, (*ii*) which interview data sources were available, (*iii*) which gears were included, and (*iv*) whether interview data needed to be shared among river sections.

'KuskoHarvEst' Solution

This issue was addressed in three primary ways. First, the code was organized in a much better fashion by "compartmentalizing" nearly every task into a designated function. This vastly reduced the amount of code needed to do repeated tasks, and functions can accept "arguments" which can be seen as settings for how it should carry out the task. Second, and supported by the first, was the development of a set of "swappable templates" of Rmarkdown code that could be readily and automatically included/excluded to complete a specific task under specific conditions. Essentially, any code that the user previously needed to edit by hand to address items like (*i*) - (*iv*) above are now contained as separate files, and the proper ones are used when the circumstances require. Third, and supported by the first two, was the construction of interactive interfaces (implemented as RStudio Add-ins via R package 'shiny', Chang et al. 2021), so the user never needs to look at or edit any code to produce an estimate. The selections the user makes in the interface triggers the execution of code in the background, which then selects the correct templates to use automatically.

Now the necessary code for a given situation is standardized, and its selection is either automated or selected via user interface inputs. This change will reduce the possibility of errors as well as the time needed to produce an estimate once the data set is finalized.

(2) Installation Difficulties

The original estimation/reporting software required that each user place many files in specific locations, know precisely where they are, and yet never edit most of them. These files contained the code that actually did the calculations for estimation (core function source code) and had to be obtained from another person who had created estimates previously, not from a centralized location. As a result, it would be possible for there to be multiple versions of the source code and eventually there could be uncertainty about the correct version to use.

'KuskoHarvEst' Solution

This solution is made possible by bundling the code into an R package. Rather than transferring the many files from user to user via USB drive or email, users simply run these two lines of code from their RStudio console (a program they also previously needed):

```
install.packages("remotes")
remotes::install_github("bstaton1/KuskoHarvEst")
```

This will install the official and most current version of the 'KuskoHarvEst' package (which contains all custom parts of the workflow that were previously disorganized) and all packages it depends on which contain more generalized R extensions. An example of an R package that 'KuskoHarvEst' depends on is the 'lubridate' package (Grolemund and Wickham 2011), which provides tools for clean and easy handling of date and time variables. After this, users only need to install a LaTeX distribution, which is a key part of the code-to-PDF document functionality and users have always needed. The source code now lives somewhere on the user's computer where they do not need to know where it is located, and can be easily updated to the most current version at anytime by running the code above again. Complete installation instructions can be found here.

(3) Subjectivity in Data Quality Decisions

In some (relatively rare) cases, the data from an interview look highly inconsistent with the other interviews or are outright impossible. An example of the former is an implied catch rate (expressed as salmon/trip) that is abnormally higher than the majority of other interviews. An example of the latter is a reported active fishing time that is longer than that of the trip as indicated by the reported start and end times. A related topic is the decision of how to share data across areas of the river when some locations have low interview coverage. Previously, there were no standardized rules for how these kinds of cases should be handled, meaning that it was possible for two analysts to obtain different results when given the same original input data.

'KuskoHarvEst' Solution

Early in the 'KuskoHarvEst' development process, a set of "decision rules" was proposed by the developer and feedback was gathered from a core group of collaborators. These decision rules established a programmed way that questionable interview data would be handled, so that specific and reoccurring issues could always be handled the same way. For the first example above, one way this was addressed was to discard the catch rate data for any interview that alters the average catch rate across all other interviews by more than $\pm$ 5%. For the second example, the software will automatically alter the active fishing time to be no longer than the reported trip duration. In both cases, the user is notified of the action the software took. For the data sharing case, a rule is now imposed that will share data from the closest area whenever there are fewer than 10 interviews in an area. These are useful examples of the automation that was added to these kinds of issues, but a complete list can be found here.

(4) Lack of Documentation

Previously, there was very little information about how to actually use the software to produce an estimation report from raw data. This was, at the time, not a terrible problem since only two people had ever used it and both were experienced R users. However, in the event that a new user needed to use the software, no method existed for them to learn it other than by asking one of these two individuals – such a case arose in 2021.

'KuskoHarvEst' Solution

Accompanying 'KuskoHarvEst' is a comprehensive set of documentation, all of which has been referenced in this document. It is now possible for a new user to learn the entire process from start to finish based solely on the information contained at the GitHub repository. Also included are example data files, showing the precise format that data must be entered in and the Staton (2018) report which contains a complete statistical description of the analytical methods. In addition to these helpful materials, each core function is documented with R's native help file system for users who wish to learn more about how the code works, which can be accessed from within RStudio by running, for example, `?KuskoHarvEst::bootstrap_harvest`.

Versioning and Hosting Platform

The development of 'KuskoHarvEst' was undertaken by tracking its history as a Git repository which was hosted as a remote repository on GitHub. The advantages to using Git/GitHub for software development are too numerous to name here, but in short, it is an approach for tracking changes to a code base. Each small change is accompanied by a complete snapshot of the software that can be returned to at any time, or copied to try out new features. Changes can be merged into the main version if they are deemed desirable or discarded entirely without ever affecting the main version. At many points in the development process, developers can comment and document why changes were necessary and what exactly was changed. One method is the issue tracker, where items that need addressing are posted. Another method is in the pull request tracker, where changes can be described to document their purpose and how they were implemented. When done properly, use of this overall strategy can lead to better software development. Although the 'KuskoHarvEst' developer is not an expert in this topic, he believes this project was one of his most well-organized and well-documented to-date. Over the development period from start (January 30, 2021) to finish (September 19, 2021), a total of 96 issues were created and addressed, 74 pull requests (groups of small changes) were opened and merged, and 457 commits (small changes) were made – together these items document the logic and implementation of development decisions.

Hosting 'KuskoHarvEst' on GitHub comes with the added benefit of a centralized official version that can be easily installed from any computer with internet access.

Future Directions

The current version of 'KuskoHarvEst' (v1.1.0) is usable indefinitely into the future without updates under two conditions. First, and most importantly, the fishery conditions and monitoring program remain should relatively consistent for its continued use. The estimators used by the software assume the data are a random sample of the fishery. This is reasonable to assume during short duration fishing opportunities with near-complete temporal coverage and decent spatial coverage of interviewers, but would become substantially more difficult in years where longer fishing opportunities are allowed (see discussion section of recent end-of-season reports for details, e.g., Russell et al. 2021). Second, 'KuskoHarvEst' depends on program R and several R packages in order to work properly. If the developers of the base R language or these additional packages fundamentally change components of their software, it could break 'KuskoHarvEst' and updates may be necessary. Given the wide usage and age of the packages that 'KuskoHarvEst' depends on, these kinds of breaking changes are unlikely.

The development of 'KuskoHarvEst' will facilitate another project (AYK-SSI project #AC-2106) that seeks to compile the historical in-season harvest monitoring data and build predictive relationships. As part of this project, all historical harvest estimates will be reproduced to ensure consistency among estimates – the standardization of previously subjective decisions and automation of many tasks in the workflow provided by 'KuskoHarvEst' will facilitate this process and make it reproducible from the raw data sources.

Acknowledgements

Several people provided valuable feedback on the functionality included in this package, in alphabetical order they are: B. Bechtol, G. Decossas, J. Esquible, D. Lowrey, J. Spaeder, K. Russell, and K. Whitworth; the vast majority of feedback was provided by B. Bechtol, K. Russell, and K. Whitworth. L. Coggins co-developed the statistical foundations of the harvest/effort estimators with B. Staton starting in 2016, with earlier work done in 2015.

'KuskoHarvEst' is completely reliant on RStudio Desktop, Rmarkdown, and Shiny to do its job, and its development was facilitated by the 'devtools' (Wickham et al. 2020) R package. The developers of these free software products are owed gratitude for making the construction of intuitive workflows like those contained in 'KuskoHarvEst' possible.

Funding for the development of 'KuskoHarvEst' was provided by the Kuskokwim River Inter-Tribal Fish Commission, administered initially by the Bering Sea Fisherman's Association through grant #AC-2101 to Quantitative Ecological Services, LLC and extended by Kuskokwim River Inter-Tribal Fish Commission grant #AC-2105 to cover some development items that took longer than originally anticipated and some additional items not covered by #AC-2101, like contributing to the final season-wide monitoring report. The initial development of the statistical and reporting framework, off of which 'KuskoHarvEst' is based, was funded by the U.S. Fish and Wildlife Service through a Pathways Position during the summers of 2016 – 2018.

References

Chang, W., Cheng, J., Allaire, J., Sievert, C., Schloerke, B., Xie, Y., Allen, J., McPherson, J., Dipert, A., and Borges, B. 2021. *shiny: Web Application Framework for R*. R package version 1.6.0; https://CRAN.R-project.org/package=shiny.

Decossas, G. 2019. In-season harvest and effort estimates for the 2019 Kuskokwim River subsistence salmon fisheries during block openers. Project summary report, Yukon Delta National Wildlife Refuge, U.S. Fish and Wildlife Service, Bethel, AK. https://www.fws.gov/uploadedFiles/Region_7/NWRS/Zone_1/Yukon_Delta/Sections/What_We_Do/Science/2019%20Subsistence%20Fisheries%20Harvest%20Report%20.pdf.

Decossas, G. 2020. In-season harvest and effort estimates for the 2020 Kuskokwim River subsistence salmon fisheries during block openers. Project summary report, Yukon Delta National Wildlife Refuge, U.S. Fish and Wildlife Service, Bethel, AK. https://www.fws.gov/uploadedFiles/2020_Fisheries_Harvest%20Report_final.pdf.

Grolemund, G. and Wickham, H. 2011. Dates and times made easy with lubridate. *Journal of Statistical Software*, 40(3):1–25. http://www.jstatsoft.org/v40/i03/.

R Core Team 2020. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. R version 4.0.2, https://www.R-project.org/.

Russell, K., Whitworth, K., Bechtol, W. R., and Staton, B. A. 2021. Inseason harvest and effort estimates for the 2021 Kuskokwim River subsistence salmon fisheries during block openers. Project summary report, Bethel, AK.

Staton, B. A. 2018. In-season harvest and effort estimates for 2018 Kuskokwim River subsistence salmon fisheries during block openers. Project summary report, Yukon Delta National Wildlife Refuge, U.S. Fish and Wildlife Service, Bethel, AK. https://www.fws.gov/uploadedFiles/2018KuskokwimRiverSalmonSubsistenceHarvestReport.pdf.

Staton, B. A. 2021. *KuskoHarvEst: Tools for Producing In-season Estimates of Salmon Harvest and Effort Occurring in Short-Duration Subsistence Harvest Opportunities in the Lower Kuskokwim River*. R package version 1.1.0, https://www.github.com/bstaton1/KuskoHarvEst.

Staton, B. A. and Coggins, Jr., L. G. 2016. In-season harvest and effort estimates for 2016 Kuskokwim River subsistence salmon fisheries during block openers. Project summary report, Yukon Delta National Wildlife Refuge, U.S. Fish and Wildlife Service, Bethel, AK. https://www.fws.gov/uploadedFiles/2016KuskokwimRiverSubsistenceSalmonHarvest.pdf.

Staton, B. A. and Coggins, Jr., L. G. 2017. In-season harvest and effort estimates for 2017 Kuskokwim River subsistence salmon fisheries during block openers. Project summary report, Yukon Delta National Wildlife Refuge, U.S. Fish and Wildlife Service, Bethel, AK. https://www.fws.gov/uploadedFiles/2017KuskokwimRiverSubsistenceSalmonHarvest.pdf.

Wickham, H., Hester, J., and Chang, W. 2020. *devtools: Tools to Make Developing R Packages Easier*. R package version 2.3.0, https://devtools.r-lib.org/.

Xie, Y. 2020. *knitr: A General-Purpose Package for Dynamic Report Generation in R*. R package version 1.29, https://yihui.org/knitr/.

Xie, Y., Allaire, J., and Grolemund, G. 2018. *R Markdown: The Definitive Guide*. Chapman and Hall/CRC, Boca Raton, Florida. https://bookdown.org/yihui/rmarkdown.

Zhu, H. 2021. *kableExtra: Construct Complex Table with 'kable' and Pipe Syntax*. R package version 1.3.4, http://haozhu233.github.io/kableExtra/,https://github.com/haozhu233/kableExtra.

Appendix A: Example In-Season Harvest Estimate Document

FIGURE A1. Example first page of the in-season harvest estimate report generated by 'KuskoHarvEst', which shows the most critical information: data source breakdowns, estimates of harvest by species/gear type, and effort by gear type. Note that because the documents are rendered from Rmarkdown source code, the analyst need not ever type any of this information into the document by hand.

Kuskokwim River In-season Harvest and Effort Estimates

6/19/2021 Subsistence Harvest Opportunity (Drift & Set Nets)

Opportunity Time Period: 6:00 AM – 6:00 PM (12 Hours)
Area Covered by Estimates: Tuntutuliak ⟷ Akiak
Contact Person(s): Kevin Whitworth (kevinwhitworth@kritfc.org), Katie Russell (krussell@nativecouncil.org)
Special Action #: 3-KS-02-21
Special Action: https://www.fws.gov/uploadedFiles/2021_fisheries_3-KS-02-21_Final_6.17.2021.pdf

Data Sources

TABLE 1. The number and percent of fisher interviews conducted by location and organization.

Data Source	Interviews	Percent
Bethel Boat Harbor (ONC)	119	50%
Other Villages (BSFA/KRITFC)	88	37%
Bethel Area Fish Camps (ONC)	30	13%
Total	**237**	**100%**

Of these interviews, **226** were from drift nets and **11** were from set nets.

TABLE 2. The time each flight was conducted and fishers counted each flight.

Time Information			Nets Counted	
Start Time	End Time	Hours	Drift	Set
9:34 AM	11:00 AM	1.43	437	12
3:13 PM	4:32 PM	1.32	271	28

Effort Estimates

- An estimated **511** total drift boat trips occurred.
  - An estimated **80%** of the trips counted on flight 2 were also counted on flight 1.
  - An estimated **20** trips were not counted during any flight.
- An estimated **31** total set net trips occurred.

Harvest Estimates

- An estimated total of **9,580 (8,400 – 10,920)** salmon were harvested.
  - An estimated total of **6,190 (5,330 – 7,140)** Chinook salmon were harvested.
  - An estimated total of **990 (670 – 1,400)** chum salmon were harvested.
  - An estimated total of **2,400 (2,080 – 2,740)** sockeye salmon were harvested.
- Harvest by set nets accounted for an estimated **630 (350 – 940)** total salmon (62% Chinook salmon, 11% chum salmon, and 27% sockeye salmon).

In-season Harvest and Effort Estimates 1 *6/19/2021 Opportunity (Drift & Set Nets)*

FIGURE A2. Example second page of the in-season harvest estimate report generated by 'Kusko-HarvEst', which shows more detailed summaries of harvest and effort by river section and the distributions of key variables across interviews.

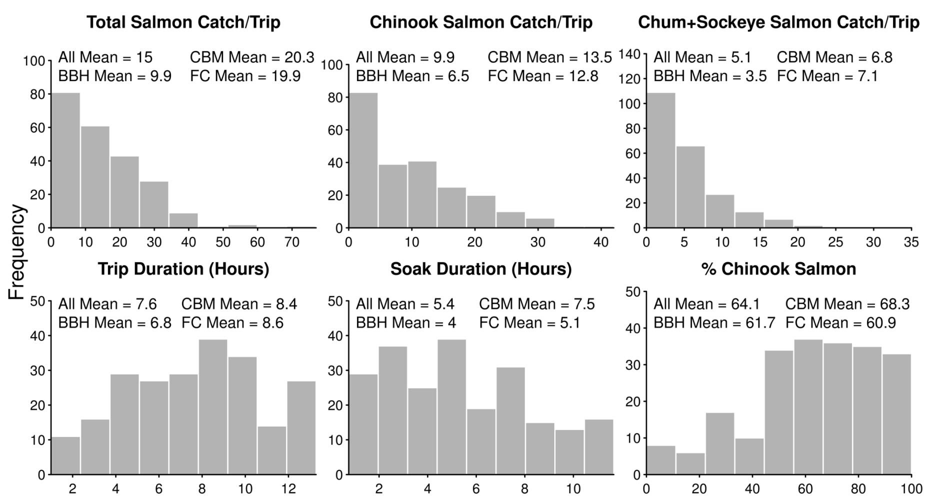
TABLE 3. Summary of relevant quantities by river stratum (area) for drift nets. Numbers in parentheses are 95% confidence intervals.

			Estimated Harvest			
Stratum	**Interviews**	**Effort Est.**	**Chinook**	**Chum**	**Sockeye**	**Total**
Tuntutuliak ⟷ Johnson R.	46	150	2,000 (1,360 – 2,790)	510 (230 – 900)	720 (490 – 970)	**3,230 (2,290 – 4,350)**
Johnson R. ⟷ Napaskiak	69	104	1,330 (1,060 – 1,640)	200 (150 – 270)	480 (370 – 580)	**2,010 (1,650 – 2,380)**
Napaskiak ⟷ Akiachak	111	212	2,040 (1,650 – 2,480)	180 (60 – 350)	860 (670 – 1,070)	**3,070 (2,500 – 3,650)**
Akiachak ⟷ Akiak	0	45	430 (340 – 520)	40 (20 – 70)	180 (140 – 220)	**650 (530 – 770)**
All	**226**	**511**	**5,810 (4,970 – 6,740)**	**920 (610 – 1,330)**	**2,230 (1,910 – 2,570)**	**8,950 (7,830 – 10,230)**

TABLE 4. Average (95% confidence limits) total salmon catch per trip and percent Chinook salmon, summarized for the areas above and below the confluence of the Johnson River with the Kuskokwim River. Quantities are derived from the strata- and species-specific harvest estimates, not the raw interview data.

	Proximity to Johnson R. Mouth	
Quantity	**Downstream**	**Upstream**
Total Catch/Trip	22 (15 – 29)	16 (14 – 18)
% Chinook Salmon	62% (55% – 68%)	66% (63% – 69%)

FIGURE 1. Distributions of relevant quantities from all completed trips using drift nets. The mean quantity by primary data source is shown in the top right; BBH = Bethel Boat Harbor (ONC), CBM = Other Villages (BSFA/KRITFC), FC = Bethel Area Fish Camps (ONC).



In-season Harvest and Effort Estimates 2 *6/19/2021 Opportunity (Drift & Set Nets)*

FIGURE A3. Example third page of the in-season harvest estimate report generated by 'KuskoHarvEst', which shows detailed summaries of the raw interview data.

Appendix: Detailed Interview Summaries

Column Meanings

- **Area**: the area of the river the trip occurred in
- **N**: the number of interviews with usable information in each area
- **Min**: the minimum value among trips in each area
- **25%**: the value that 25% of trips fell below in each area
- **Mean**: the average value across trips in each area
- **75%**: the value that 75% of trips fell below in each area
- **Max**: the maximum value among trips in each area

Information is for drift net trips only.

TABLE A1. Summary of drift net catch rate of Chinook salmon by fishing area (salmon per 150 feet of net per hour).

Area	N	Min	25%	Mean	75%	Max
Tuntutuliak ⟷ Johnson R.	46	0	0.3	1.1	1.4	9.6
Johnson R. ⟷ Napaskiak	69	0	0.9	2.2	2.8	9.8
Napaskiak ⟷ Akiachak	111	0	0.7	2.2	2.9	14.5
All	**226**	**0**	**0.6**	**2**	**2.6**	**14.5**

TABLE A2. Summary of drift net catch per trip of Chinook salmon by fishing area.

Area	N	Min	25%	Mean	75%	Max
Tuntutuliak ⟷ Johnson R.	46	0	4	10	13	42
Johnson R. ⟷ Napaskiak	69	0	4	11	17	31
Napaskiak ⟷ Akiachak	111	0	2	9	14	32
All	**226**	**0**	**3**	**10**	**15**	**42**

TABLE A3. Summary of drift net catch rate of chum+sockeye salmon by fishing area (salmon per 150 feet of net per hour).

Area	N	Min	25%	Mean	75%	Max
Tuntutuliak ⟷ Johnson R.	46	0	0.2	0.7	0.8	4.8
Johnson R. ⟷ Napaskiak	69	0	0.3	1.1	1.7	3.6
Napaskiak ⟷ Akiachak	111	0	0.3	1.1	1.4	8.4
All	**226**	**0**	**0.3**	**1**	**1.4**	**8.4**

TABLE A4. Summary of drift net catch per trip of chum+sockeye salmon by fishing area.

Area	N	Min	25%	Mean	75%	Max
Tuntutuliak ⟷ Johnson R.	46	0	2	6	7	35
Johnson R. ⟷ Napaskiak	69	0	2	5	8	16
Napaskiak ⟷ Akiachak	111	0	1	5	6	31
All	**226**	**0**	**2**	**5**	**7**	**35**

In-season Harvest and Effort Estimates 3 *6/19/2021 Opportunity (Drift & Set Nets)*

FIGURE A4. Example third page of the in-season harvest estimate report generated by 'KuskoHarvEst', which shows additional detailed summaries of the raw interview data.

TABLE A5. Summary of drift net percent composition of Chinook salmon by fishing area.

Area	N	Min	25%	Mean	75%	Max
Tuntutuliak ⟷ Johnson R.	46	0%	50%	60%	77%	100%
Johnson R. ⟷ Napaskiak	69	0%	55%	66%	82%	100%
Napaskiak ⟷ Akiachak	111	0%	50%	65%	83%	100%
All	**226**	**0%**	**50%**	**64%**	**81%**	**100%**

TABLE A6. Summary of drift net active fishing hours by fishing area.

Area	N	Min	25%	Mean	75%	Max
Tuntutuliak ⟷ Johnson R.	45	0.8	3.8	6.5	10	11.2
Johnson R. ⟷ Napaskiak	69	1	3.5	6.1	8.5	11.5
Napaskiak ⟷ Akiachak	110	0.8	3	4.5	6	11.7
All	**224**	**0.8**	**3**	**5.4**	**7.5**	**11.7**

TABLE A7. Summary of drift net total trip duration by fishing area.

Area	N	Min	25%	Mean	75%	Max
Tuntutuliak ⟷ Johnson R.	46	1.5	6.3	8.7	10.9	12.8
Johnson R. ⟷ Napaskiak	69	1.9	5.5	7.8	9.5	12.9
Napaskiak ⟷ Akiachak	111	1	4.3	7	9.4	13.3
All	**226**	**1**	**5.2**	**7.6**	**10**	**13.3**

TABLE A8. Summary of drift net trip start time by fishing area.

Area	N	Min	25%	Mean	75%	Max
Tuntutuliak ⟷ Johnson R.	46	5:00 AM	6:00 AM	7:41 AM	9:30 AM	12:00 PM
Johnson R. ⟷ Napaskiak	69	5:30 AM	6:00 AM	8:07 AM	9:30 AM	3:30 PM
Napaskiak ⟷ Akiachak	111	5:00 AM	6:00 AM	8:15 AM	10:00 AM	4:00 PM
All	**226**	**5:00 AM**	**6:00 AM**	**8:06 AM**	**9:30 AM**	**4:00 PM**

TABLE A9. Summary of drift net trip end time by fishing area.

Area	N	Min	25%	Mean	75%	Max
Tuntutuliak ⟷ Johnson R.	46	8:22 AM	3:00 PM	4:21 PM	6:00 PM	7:30 PM
Johnson R. ⟷ Napaskiak	69	11:00 AM	2:03 PM	3:58 PM	5:45 PM	7:57 PM
Napaskiak ⟷ Akiachak	111	8:48 AM	1:00 PM	3:16 PM	5:30 PM	7:30 PM
All	**226**	**8:22 AM**	**2:00 PM**	**3:42 PM**	**6:00 PM**	**7:57 PM**

Appendix B: Example Sensitivity Analysis Document

FIGURE B1. Example first page of the sensitivity analysis report generated by 'KuskoHarvEst', which shows the results of effort sensitivity analyses where some of the aerial survey data were excluded.

Kuskokwim River In-season Harvest/Effort Sensitivity Analysis

6/19/2021 Subsistence Harvest Opportunity (Drift & Set Nets)

Opportunity Time Period: 6:00 AM – 6:00 PM (12 Hours)
Area Covered by Estimates: Tuntutuliak ⟷ Akiak
Contact Person(s): Kevin Whitworth (kevinwhitworth@kritfc.org), Katie Russell (krussell@nativecouncil.org)
Special Action #: 3-KS-02-21
Special Action: https://www.fws.gov/uploadedFiles/2021_fisheries_3-KS-02-21_Final_6.17.2021.pdf

Effort Sensitivity Analyses

In the tables below, "Pr(F1|F2)" is the probability that a trip counted on flight 2 was also counted on flight 1 (the estimator corrects for double-counting; this column will not display if only one flight was made). "Trips per interview" is an estimate of how many actual trips occurred per interview that was conducted, and is used to expand the effort estimate for interviews that are known to have occurred but could not have been counted because they started and stopped during times when the flights were inactive. "Trips not observed" is the estimated number of trips that fall in this expanded category.

Exclusion of Flight Data

This analysis leaves out all combinations of flight data and re-calculates the drift net effort estimates. This helps give a sense of how the estimate would have been different had fewer flights been conducted. Generally, you want to see small changes because with good interview data coverage, the effort estimator should be robust to fewer flights. If there are large changes (e.g., >50%), that indicates that the distribution of fishing effort indicated by the interview data and the flight data are inconsistent, and emphasizes that conducting several flights within an opportunity is important to obtain a good estimate.

Scenario	Effort Estimate	% Change	Pr(F1|F2)	Trips Per Interview	Trips Not Observed
F1, F2	511	0%	80%	2.26	20
F1	523	2%	–	2.31	86
F2	437	-14%	–	1.94	166

Sensitivity Analyses 1 *6/19/2021 Opportunity (Drift & Set Nets)*

FIGURE B2. Example second page of the sensitivity analysis report generated by 'KuskoHarvEst', which shows the results of effort sensitivity analyses where some of the interview data were excluded.

Exclusion of Interview Data

This analysis leaves out all combinations of interview data sources and re-calculates the drift net effort estimates (the scenario column indicates which data sources were kept). This helps give a sense of how much each data source affects the overall estimate, and how much the estimate would have been different had fewer data sources been available. Generally, you want to see relatively small changes when leaving out data – this is an indication that each source took a random sample of the fishers it could have sampled with respect to trip times. If large changes occur (e.g., >50%), that is not necessarily a problem, but it is good information to know.

Scenario	Effort Estimate	% Change	Pr(F1|F2)	Trips Per Interview	Trips Not Observed
BBH, CBM, FC	511	0%	80%	2.26	20
BBH, CBM	517	1%	78%	2.61	21
BBH, FC	538	5%	72%	3.79	27
CBM, FC	474	-7%	91%	4.23	13
BBH	558	9%	66%	4.90	29
CBM	474	-7%	90%	5.66	11
FC	469	-8%	94%	16.78	17

Sensitivity Analyses 2 *6/19/2021 Opportunity (Drift & Set Nets)*

FIGURE B3. Example third page of the sensitivity analysis report generated by 'KuskoHarvEst', which shows the results of harvest sensitivity analyses where some of the interview data were excluded.

Harvest Sensitivity Analyses

This analysis leaves out one interview data source at a time and recalculates the harvest estimates for all species and gears. The effort estimate still uses all data. This gives information about how much each data source contributes to the overall estimates, and how different the estimates would have been had a data source not been collected. The "Estimate" column shows the mean estimate and 95% confidence interval in parentheses, the "% Change" column gives a measure of change relative to the mean estimate using all data, and the "CV" column is the coefficient of variation of the estimate – this is a measure of uncertainty (higher values mean more uncertainty).

	Chinook			Chum			Sockeye			Total		
Scenario	**Estimate**	**% Change**	**CV**	**Estimate**	**% Change**	**CV**	**Estimate**	**% Change**	**CV**	**Estimate**	**% Change**	**CV**
All Data	6,190 (5,330 – 7,140)	0%	8%	990 (670 – 1,400)	0%	19%	2,400 (2,080 – 2,740)	0%	7%	9,580 (8,400 – 10,920)	0%	7%
No FC	6,060 (5,080 – 7,200)	-2%	9%	1,040 (660 – 1,550)	5%	22%	2,210 (1,860 – 2,550)	-8%	8%	9,310 (8,020 – 10,860)	-3%	8%
No CBM	4,590 (3,700 – 5,590)	-26%	10%	640 (400 – 920)	-35%	20%	2,130 (1,790 – 2,490)	-11%	8%	7,360 (6,210 – 8,660)	-23%	8%
No BBH	8,620 (7,360 – 9,990)	39%	8%	1,150 (730 – 1,780)	16%	24%	3,270 (2,700 – 3,880)	36%	10%	13,030 (11,260 – 15,070)	36%	7%

Sensitivity Analyses 3 *6/19/2021 Opportunity (Drift & Set Nets)*