SUPPORTING STATEMENT B FOR PAPERWORK REDUCTION ACT SUBMISSION

In-Season Subsistence Salmon Fishery Catch and Effort Survey OMB Control Number 1018-NEW

1. Describe (including a numerical estimate) the potential respondent universe and any sampling or other respondent selection method to be used. Data on the number of entities (e.g., establishments, State and local government units, households, or persons) in the universe covered by the collection and in the corresponding sample are to be provided in tabular form for the universe as a whole and for each of the strata in the proposed sample. Indicate expected response rates for the collection as a whole. If the collection had been conducted previously, include the actual response rate achieved during the last collection.

The potential respondent universe is the population of fishers that harvest salmon during a subsistence fishery opening. In 2018 (the last time these surveys were conducted), the estimated number of fisher trips ranged between 276 and 466. In the openings corresponding to these fisher estimates, 29 and 97 interviews were conducted at one location, the Bethel Boat Harbor, respectively. This represents an estimated response rate of approximately 10% - 20%. Not all of the estimated fishers were available for sampling by interviewers at the Bethel Boat Harbor, given there are many other access points to the fishery. Not all individuals agree to complete the survey and there is no information on refusal rates by the tribal organization conducting this survey. Other projects are currently in place that sample these locations that use the same survey. In 2018, the average total sampling rate across all sites was approximately 50%.

2. Describe the procedures for the collection of information including:
- *** Statistical methodology for stratification and sample selection,**
- *** Estimation procedure,**
- *** Degree of accuracy needed for the purpose described in the justification,**
- *** Unusual problems requiring specialized sampling procedures, and**
- *** Any use of periodic (less frequent than annual) data collection cycles to reduce burden.**

The in-season harvest estimation framework that was developed and applied to the 2016 – 2018 Kuskokwim River salmon seasons required two primary types of information:

(1) An estimate of the total number of fishing trips each day; and
(2) Completed trip interview information from fishers documenting gear, fishing location, fishing time, and catch.

The methods used to estimate harvest in 2018 were identical to those used in 2017, except for one additional location (see Section 2.2).

2.1 Aerial Net Counts

For each opener, which is described as a windowed fishing opportunity to harvest salmon, two or more aerial survey flights were flown to count the number of drift boats and set nets fishing

within the YDNWR boundaries (Figure 1). Flights were scheduled to capture boat counts between low and high tide when the tides were moving the strongest, which are the most popular times to fish; the flights were spaced relatively equally throughout the opener. Oftentimes, this resulted in approximately 3 – 4 hours between the end of one flight and the start of the next flight (Table 1). Flight missions involved a USFWS pilot and at least one observer flying at an altitude of 500 – 700ft and using predominately YDNWR aircraft (most frequently Cessna 185 n714 and Cessna 2016 n740).

Flight missions involved departing Bethel, following the river downstream and southwest toward Kuskokwim Bay to the village of Tuntutuliak, then turning upstream and northeast to fly to the village of Akiak by following the river (Figure 1). This path took approximately 1.5 hours to complete, including the flight back to the Bethel airport. All sections of the flight path along the main stem Kuskokwim River, excluding below Loumavik Slough (which is just downstream of the Johnson River confluence) and Kuskokuak Slough (off of which the Kwethluk and Kisaralik Rivers branch), were counted twice (i.e., once flying downstream, once flying upstream). The maximum of the two counts was used as the boat count for that section. Below Loumavik Slough, the river is too wide to see both banks entirely so each bank was counted once and the counts were summed. The very small amount of nets (typically <5) observed in Kuskokuak Slough were not included in the count given the coverage of the harvest estimation was limited to the main stem Kuskokwim River only.

Drift boat and set net counts were recorded into approximately 10 river regions demarcated by major landmarks (e.g., villages or tributaries) and then assigned to four strata (Figure 1, denoted by the letters A – D). Boats were counted if they were actively fishing or if a net could be seen in the boat. If there was any doubt about whether boats were fishing boats (such as below Loumavik Slough, where large distances made it difficult to look for nets inside boats in transit), they were counted as fishing boats. On two occasions, inclement weather prevented USFWS from flying scheduled effort surveys: 6/16/2018 and 7/5/2018. Both of these canceled flights were the first of three scheduled flights for 12-hour openers, though the latter two flights each day were flown with no issues.

2.2 Completed Trip Interviews

Information from fisher trips was obtained from the following five sources:

(1) the Bethel boat harbor;
(2) Bethel area fish camps;
(3) several main stem villages other than Bethel;
(4) several tundra villages; and
(5) from USFWS law enforcement personnel during routine roving compliance checks.

Interview data from sources (1) and (2) were collected by personnel from the Orutsararmiut Traditional Native Council (ONC) and were the predominate sources used by Staton and Coggins (2016). Data from source (3) were collected beginning in 2017 as part of a community-based monitoring project established by the Bering Sea Fisherman's Association (BSFA) to, among other things, provide interview data from areas of the YDNWR other than solely the Bethel area. The BSFA village monitors in 2018 were located in the villages of Tuntutuliak, Napaskiak, Kwethluk, Akiachak, and Akiak and reported data in a timely manner so that they could be included into the estimates. Data from source (4) were collected by ADF&G Division of Subsistence staff stationed in the tundra villages of Atmautluak and Kasigluk — data from this source were a new addition for 2018. Data from source (5) have been available since 2015, but

have been of varying quality. The data from source (5) collected in 2018 were very high quality and were included when available (all openers except the final one on 7/5/2018 before restrictions were lifted). It should be noted, however, that because the law enforcement interviews were not completed trips, the only information that was used from these interviews was the catch rate and the net length (see Section 2.3.3).

Interviewees sampled by these five sources were asked the same questions and the interviewers were trained as thoroughly as possible in a formal setting (in Bethel; 5/29 – 5/30 of 2018) to ensure the questions were asked in a consistent fashion. Interviewers were instructed to spend as much time as possible collecting data during openers, which during the short openers (none longer than 12 hours) allowed for nearly complete coverage at interview locations. Interviews were intended to be minimally intrusive yet still gain accurate and meaningful information regarding the fishing trip. The key pieces of information collected in each interview (indexed by *i*) included:

- The day fishing occurred (indexed by *d*)
- The location of the trip (used to place the trip in a geographic stratum, indexed by *j*; Figure 1)
- The type of net used (drift versus set net)
- The start and end times of the trip ($T_{1,i,d}$ and $T_{2,i,d}$, respectively)
- The total number of hours the net was fishing (referred to as "soak time"; $h_{i,j,d}$)
- The length of the net used (in feet; $L_{i,j,d}$)
- The total harvest by species of each Chinook, chum, and sockeye salmon ($C_{i,j,s,d}$; species indexed by *s*)

2.3 Analytical Methods

Although the analytical methods used in 2018 were nearly identical to those presented in Staton and Coggins (2016), a full description will be provided here for completeness.

2.3.1 Boat Effort Expansion Model

When interpreting aerial survey counts, it is important to consider two facts which result from the counts being instantaneous surveys rather than complete censuses. First, some active drift boat trips counted during one flight were likely also active in subsequent flights (i.e., some boats were double- or triple-counted). Second, surely some number of drift boat trips started and ended during times that were not flown (i.e., some boats fished but were not counted). Thus, to obtain an estimate of the total number of drift boat trips in an opener, a method was needed to correct for these two issues. It is possible to derive an estimator for the number of boat trips on day *d* based on:

(1) the boat counts made on each flight $c(A_{c,d})$,
(2) the start ($F_{1,c,d}$) and stop ($F_{2,c,d}$) times of flight *c* on the same day, and
(3) the start ($T_{1,i,d}$) and stop ($T_{2,d}$) times of each of the n_d completed trip interviews.

The estimator operates by determining if each interviewed trip *i* was actively fishing during a period when it could have been counted on flight *c* using numerical logic (i.e., Boolean operators). For example, trip *i* would have been counted on flight *c* if the trip started before the flight started and ended after the flight ended (i.e., if both conditions $T_{1,i,d} < F_{1,c,d}$ and $T_{2,i,d} >$

$F_{2,c,d}$ were met). This can be expressed more simply by determining if each trip was not available to be counted, i.e., if it started and ended either before or after the flight:

$$T_{1,i,d} < T_{2,i,d} < F_{1,c,d}, \qquad (1)$$

$$F_{2,c,d} < T_{1,i,d} < T_{2,i,d}. \qquad (2)$$

If either of the conditions (1) or (2) is met, trip *i* was not available to be counted *via* aerial survey flight *c*. These conditions were tested for each flight *c* for each of the n_d interviews and the following summaries were calculated:

- The number of interviewed trips available to be counted on flight *c* ($B_{c,d}$)
- The number of interviewed trips available to be counted on two consecutive flights *c* and *c* **+ 1** ($B_{c,c+1,d}$)
- The number of interviewed trips available to be counted during at least one flight ($B_{y,d}$)
- The number of interviewed trips not available to be counted during any of the flights that occurred on day *d* ($B_{n,d}$)

Based on these quantities, the effort expansion model corrected each aerial count ($A_{c,d}$) by the number of trips counted on the previous flight of the day. First, the proportion of interviewed boat trips counted on flight *c* **+ 1** that were also counted on flight c (denoted $p_{old,c}$) was calculated:

$$p_{old,c} = \frac{B_{c,c+1,d}}{B_{c+1,d}}. \qquad (3)$$

which is an estimator of the joint probability of a boat being counted on two consecutive flights *c* and *c* **+ 1**. The quantity $p_{old,c}$ should be a function of *(a)* the magnitude of $F_{1,c+1,d} - F_{2,c,d}$ and *(b)* the magnitude of the average $T_{2,i,d}$ *and* $T_{1,i,d}$ (openers with closely-spaced flights and long trips will have higher values of $p_{old,c}$, i.e., rates of double-counting). The number of trips that were counted on flight *c* **+ 1** that were not counted on flight *c* is then:

$$\hat{A}_{c+1,d} = A_{c+1,d}(1 - p_{old,c}). \qquad (4)$$

Note that Equation (4) need only be calculated for flights *c* > 1, as all trips counted on flight *c* **= 1** were new entries to the fishery as far as the estimator is concerned. The total number of boat trips that were counted during a flight is:

$$\hat{B}_d = A_{1,d} + \sum_{z=2} \hat{A}_{z,d}. \qquad (5)$$

To correct the count for trips that occurred between flights (it is known that at least $B_{n,d}$ such trips occurred), a simple scaling method based on the Petersen estimator (Seber, 1982) was applied and the result was added to $\hat{B}_d$ to obtain the total number of drift trips during the opener:

$$\hat{B}_d = B_{n,d} \frac{\hat{B}_d}{B_{y,d}} + \hat{B}_d. \qquad (6)$$

Given that the sampling plan for interviews did not involve stratification by location, this estimator was applied by aggregating all completed trip interviews regardless of stratum. The total estimated drift boat trips ($\hat{B}_d$) was post-stratified into geographic strata based on the average proportion of boats counted in each stratum across all flights that day:

$$p_{j,d} = \frac{\sum_{c} A_{c,j,d}}{\sum_{cj} A_{c,j,d}}. \qquad (7)$$

And

$$\hat{B}_{j,d} = p_{j,d}\hat{B}_d. \qquad (8)$$

2.3.2 Set Net Effort Expansion Model

Due to a severe lack of interviews from set net fishers, it was not possible to use the procedure described above for drift boat fishers for set nets. To account for daily set net effort, the sum of the maximum set net aerial count from each geographic stratum was used as the effort for that day.

2.3.3 Harvest Expansion Model

The harvest expansion model used the two pieces of information (trip characteristics from interviews and total effort estimates) to estimate the total harvest by geographic stratum and opener. An index of trip-specific effort was obtained:

$$e_{i,j,d} = L_{i,j,d}h_{i,j,d}, \qquad (9)$$

where the units of trip-level effort are in net-foot-hours, and was used to account for the observation that fishers use different lengths of net. Then, a catch rate was calculated for each species to standardize catch numbers across trips:

$$\mathrm{CPUE}_{i,j,s,d} = \frac{C_{i,j,s,d}}{e_{i,j,d}}. \qquad (10)$$

The quantities $L_{i,j,d}$, $h_{i,j,d}$, and $CPUE_{i,j,s,d}$ were averaged across all interviews available in a geographic stratum to obtain the characteristics of the average trip occurring on day d in geographic stratum j ($L_{j,d}$, $h_{j,d}$, and $CPUE_{j,s,d}$). Total estimated harvest of species s for stratum j on day d is then:

$$\hat{H}_{j,s,d} = \hat{B}_{i,d}\bar{L}_{i,d}\bar{h}_{i,d}\mathrm{CPUE}_{j,s,d}. \qquad (11)$$

This expansion was conducted separately for drift net fishers and set net fishers, using only the interview data from each gear type. It was conducted as geographically-explicitly as possible. As a general rule, if a domain had fewer than 15 interviews, interview data from the nearest stratum were were aggregated with the data-poor stratum when calculating the average quantities $L_{j,d}$, $h_{j,d}$, and $CPUE_{j,s,d}$.

Total harvest by species on day d was calculated by summing the strata-specific estimates:

$$\hat{H}_{s,d} = \sum_{j} \hat{H}_{j,s,d}, \qquad (12)$$

and total salmon harvest for day *d* was calculated by summing across species-specific estimates:

$$\hat{H}_d = \sum_s \hat{H}_{s,d} \qquad (13)$$

2.3.4 Uncertainty Estimation

Variability in among-interview quantities was quite high (particularly for $CPUE_{i,j,s,d}$ and $h_{i,j,d}$), necessitating the consideration of statistical uncertainty in the estimates. Uncertainty was quantified using a non-parametric bootstrap. Bootstrapping involves randomly sampling (with replacement) from the observed trip interviews, producing a harvest expansion estimate following the above method for each randomized data set (Equations 9 – 13), and repeating the process many times (10,000 in this case) to form a distribution of possible harvests given the observed sample of interviews. To summarize the resulting variation, the 2.5th and 97.5th quantiles were used as the lower and upper confidence limits (CL), respectively, and the mean of all bootstrapped estimates was used as the point (i.e., most likely) estimate.

While there are other methods to estimate uncertainty in the harvest estimates, it was determined that the non-parametric bootstrap was the most appropriate method because other methods make a variety of tenuous assumptions (Efron and Tibshirani, 1993). It is important to recognize that the harvest estimates contained in this report do not account for sampling variability in the process of estimating effort (i.e., boat trips; Equations 3 – 8) during aerial surveys. Thus, uncertainty in the harvest estimates is smaller than if uncertainty in effort was fully considered.

3. Describe methods to maximize response rates and to deal with issues of non-response. The accuracy and reliability of information collected must be shown to be adequate for intended uses. For collections based on sampling, a special justification must be provided for any collection that will not yield "reliable" data that can be generalized to the universe studied.

Given the data collection is an on-site in-person interview, the response rate is influenced by how many sites are sampled and how long interviewers are stationed at each site. Previous sampling programs have targeted the Bethel Boat Harbor because it is the busiest known fishery access point. Whenever possible, interviewers have been stationed there for the entire time that fishers (potential respondents) may return and be able to be interviewed. This is highly burdensome on the organizations collecting the data, as they are tasked with sampling at other sites as well.

This Information Collection request is to support additional data collection by FWS at the Bethel Boat Harbor. Thus, the purpose of this request is to help maximize the response rate of the survey. Non-response occurs at a sampled site for two reasons: either an interviewer is unavailable to interview the fisher or the fisher declines to be interviewed (the surveys are voluntary). Placing more interviewers, as is the purpose of this request, can help reduce the frequency of occurrence for the former cause of non-response. The accuracy and reliability of the information collected has been verified in previous years using several methods.

First, if there is any inconsistency in the information collected in a single interview, that interview would be discarded (this has been very rare – USFWS estimate less than 5% of all interviews

conducted). Second, the organizations that have collected the data in the past have had a chance to review the way the data were used and have always verified that the data are reliable with respect to their intended use (estimation of harvest). Third, post-season estimates of harvest are produced based on an independent sampling program implemented by the state of Alaska. In all years in which harvest estimates have been produced in-season based on the information collected in these interviews, the estimates have agreed very closely between these two programs. This indicates the information collected in-season is reliable enough for the purposes of harvest estimation – if it were unreliable then we would see drastic differences between these two estimation programs

4. Describe any tests of procedures or methods to be undertaken. Testing is encouraged as an effective means of refining collections of information to minimize burden and improve utility. Tests must be approved if they call for answers to identical questions from 10 or more respondents. A proposed test or set of tests may be submitted for approval separately or in combination with the main collection of information.

There are no tests of procedures or methods.

5. Provide the names and telephone numbers of individuals consulted on statistical aspects of the design and the name of the agency unit, contractor(s), grantee(s), or other person(s) who will actually collect and/or analyze the information for the agency.

1. Lewis Coggins, Yukon Delta National Wildlife Refuge, 907-543-1022
2. Ben Staton, Auburn University, Bas0041@tigermail.auburn.edu
3. Janessa Esquible, Orutsararmiut Native Council 907-543-2608
4. Gary Decossas, Yukon Delta NWR, 907-543-3151