# Alaska Statewide and Regional Estimates of Consumption Rates in Rural Communities for Salmon, Halibut, Herring, Non-Marine fish, and Marine Invertebrates

Final Report

Prepared by:
Nayak Polissar, PhD
Moni Neradilek, MS
The Mountain-Whisper-Light Statistics

Through a subcontract to
Tetra Tech
EPA Contract EP-C-14-016

March 20, 2019

# Contents

# Figures

# Tables

# Foreword

*This foreword has been authored by the U.S. Environmental Protection Agency (EPA) and Alaska Department of Environmental Conservation (ADEC), while all other sections of this report have been authored by contractors to the EPA listed on the title page.*

The State of Alaska is working to develop new and revised ambient water quality criteria to protect human health (human health criteria, or HHC). Derivation of HHC is based upon several input variables, one of which is the fish consumption rate (FCR).

Currently, there are limited Alaska fish consumption surveys that are appropriately designed and that provide sufficient information to derive statewide FCRs to support HHC development for Alaska.[1] However, the Alaska Department of Fish and Game Division of Subsistence (herein, ADF&G) has community harvest surveys that contain a wealth of information on harvest of fish and game by rural Alaskan communities. Although the community harvest surveys differ in objective from surveys designed to characterize FCRs, the harvest surveys are well designed and conducted. The harvest surveys provide statewide geographic coverage of Alaska and include temporal harvest trends. EPA and ADEC support consideration of these community harvest survey resource use rates to characterize FCRs for HHC development.

The purpose of this report is to review the ADF&G community harvest datasets, methodology for characterizing FCRs from the survey resource use rates, and analysis protocols in an effort to ensure ADEC's efforts to adopt a revised fish consumption rate are technically defensible. The report also identifies differences in community harvest and fish consumption survey methodologies and the potential impact of these differences. While it is not currently possible to fully characterize the impact of these differences, comparison of ADF&G resource use rates with FCRs from other surveys of Native American fish consumption elsewhere in EPA Region 10 states indicate that ADF&G resource use rates and locally derived FCRs are similar in magnitude. This provides further support for application of resource use rates to characterize FCRs for HHC development.

The following report concludes nearly one and a half years of work to evaluate and improve the applicability of ADF&G resource use rates to characterize Alaska FCRs for HHC development. The report provides an analysis of ADF&G fish resource use rates to ensure that they are as accurate and representative as possible for various Alaska regions and rural Alaska as a whole. Alaska resource use rates fall within the range of FCRs for other fish consuming tribes within EPA Region 10, offering further support for employing resource use rates to represent FCRs for HHC development. The information used is the best available to represent relevant fisheries resource consumption of the rural population of Alaska and to characterize FCRs for ambient water quality criteria development.

---

[1] Two surveys of Alaska Native tribal fish consumption have been completed. The first evaluates fish consumption for several villages in Cook Inlet (Seldovia Village Tribe 2013, http://svt.org/wp-content/uploads/2016/03/assessment-of-cook-inlet-tribes-subsistence-consumption-final.pdf). The second report, by the Sun'aq Tribe of Kodiak, in final draft form, evaluates fish consumption for several villages on Kodiak Island (Lance et al. 2019, http://sunaq.org/wp-content/uploads/2016/09/Kodiak-Tribes-Seafood-Consumption-Assessment-DRAFT-Final-Report-26Feb19-FINAL.pdf).

## Acknowledgements

## Executive Summary

The Alaska Department of Environmental Conservation (ADEC) requires estimates of fish and shellfish consumption rates (hereafter referred to as FCRs) for the population of Alaska to derive water quality criteria for the protection of human health. Alaska has a comprehensive database of community harvest surveys that have been conducted by the Alaska Department of Fish and Game Division of Subsistence (herein, ADF&G). ADF&G reported a reliable methodology for its field work and data collection, which is a strength supporting the survey data.

ADF&G analyzed data from its community harvest surveys conducted to derive fish and shellfish resource use rates[2] using a methodology that is designed to accommodate resource use data collected per household (Wolfe and Utermohle 2000). The surveys were conducted during 2009–2016, representing study years 2008–2015. The purpose of the current report is to re-compute ADF&G resource use rate estimates to better support ADEC's efforts to revise Alaska's criteria to protect human health.

This report presents estimated mean and 90th percentile resource use rates for a population of 262 rural census-designated communities in Alaska. Based on these 262 communities, rates have been calculated from data provided by ADF&G for a non-random sample of 110 ADF&G-defined rural study communities.[3] These use rates include the resources of salmon, halibut, herring, non-marine fish, and marine invertebrates; see Appendix A, Table A1 for a complete list of individual resources that fall under this rubric.[4] Selected rates (mean and 90th percentile) are presented for *consumers only* and for *consumers and non-consumers combined* (a per capita rate).[5] The rates are presented for each of Alaska's six regions and for all regions combined (the entire State). The geographic distribution of the rural communities is mapped in Figure C1 (showing sampled and nonsampled rural communities), and a map of Alaska's six regions is shown in Figure C2 of Appendix C.

The Mountain-Whisper-Light Statistics (TMWL) developed and used statistical weighting to adjust the non-random sample data to estimate rural use rates for the State as a whole and for each of Alaska's six regions (Arctic, Interior, Western, Southwestern, Southcentral, and Southeast). The rates presented here more accurately represent Alaska's rural fish and shellfish-consuming population than rates developed without appropriate statistical weighting. The development and use of the statistical weighting is a very important feature of this analysis. The weighting attempts to adjust for the non-random selection of communities and reduce potential bias of the estimated consumption rates for each of the six regions and for the State as a whole.

---

[2] In this report the term "resource" can be read as "species." "Use" in this report refers to human consumption.

[3] See section 2.2 "Data Used" for the description of "ADF&G-defined study communities" and their relation to U.S. census-designated communities for this report. Unless explicitly stated (such as in the statement attached to this is footnote), the term "community" in this report refers to a U.S. census-designated community.

[4] Table A1 lists resources which, generically, fall under the rubric "salmon, halibut, herring, non-marine fish, marine invertebrates." Not all of the resources listed in Table A1 were actually reported in the survey data as consumed by at least one respondent household. The resources that actually occurred in the survey data are marked by asterisks (*) in the table.

[5] People reporting zero consumption of fish and shellfish are not included in the calculation of the "consumers only" rates presented in this report. The "per capita" rates refer to the combination of both consumers and non-consumers. All individuals, including those with zero consumption, are included in the calculation of the "per capita" rate.

After applying the statistical weighting, the statewide mean *consumers only* use rate in rural communities is 149 grams per day, and the per capita mean rate (consumers and non-consumers combined) is 141 grams per day. The consumers only 90th percentile rate is 308 grams per day, and the per capita 90th percentile rate is 302 grams per day. The six regions varied widely in their use rates. For example, the Western region had the highest means and 90th percentiles, exceeding those of the Southcentral region, which had rates that were 68% to 80% lower than corresponding Western rates. The consumers only mean varied from 113 to 190 grams per day across regions (per capita range: 105–183 grams per day). The regional variation in 90th percentile consumers only rates was 217–379 grams per day (per capita range: 209–376 grams per day).

TMWL identified three sources of uncertainty in this report, the impact of which is discussed later in the report (see section 3). Despite these uncertainties, the rates estimated herein provide valuable estimates for the State to use in its regulatory decisions. The statistical methods used by TMWL utilize the data as collected and provide a picture of regional differences in resource use.

This report finds that the differences in calculated resource use rates among Alaska's six regions are not statistically significant and the margin of error of the calculated rates is relatively large. The margin of error is driven primarily by sample size and variability of consumption across households and communities. Given the margin of error and considering the large differences in calculated use rates from region to region, there may be substantial true differences in use rates among the regions. The potential sources of bias in the survey and the limitations of the survey— some are common to any survey—cannot be fully quantified. Comparison of fisheries resource use rates with subsistence fish consumption rates obtained via fish consumption surveys are of similar magnitude, further supporting characterizing FCRs using resource use rates. Resource use rates for consumers only vs. consumers and non-consumers (combined) are quite similar due to the relatively small fraction of non-consumers encountered in the surveys. In conclusion, despite limitations, the information presented in this report can be used to represent relevant fisheries resource consumption of the rural population of Alaska and to characterize FCRs for ambient water quality criteria development. The surveys clearly represent a useful gain in knowledge about resource use in rural Alaska.

# 1   Introduction

Alaska Department of Environmental Conservation (ADEC) plans to adopt new and revised human health criteria (HHC) into the state's water quality standards. These HHC represent specific levels of chemicals or conditions in a waterbody that are not expected to cause adverse effects to human health. One of the primary challenges of setting HHC is determining an appropriate fish consumption rate (FCR) based on available data (both historic and current). Of interest are rural FCRs for the entire state, the FCR for each of Alaska's six regions as defined by Alaska Department of Fish and Game Division of Subsistence (herein, ADF&G), and potential differences in FCRs among those regions.

ADEC has been collaborating with ADF&G with the goal of obtaining resource rates that are relevant for environmental regulation. ADF&G has long maintained a library of technical papers that contain information on traditional and subsistence uses of many natural resources, including harvest and use patterns through household surveys. In 2017, ADF&G completed a preliminary draft analysis of resource harvest data collected between 2009 and 2016 for rural Alaska populations (ADF&G 2017a).

The Mountain-Whisper-Light (TMWL) was retained by Tetra Tech on behalf of the U.S. Environmental Protection Agency (EPA) to review the preliminary analysis on resource use rates in the draft report prepared by ADF&G (ADF&G 2017a). The review was intended to evaluate the analytic methodology used by ADF&G in its calculation of resource use rates and ADF&G's comparison of rates among regions. This report (based on a dataset that has been revised by ADF&G) summarizes the results of TMWL's review and outlines TMWL's revised methodology and rate calculations using the new data. These changes will better support the resulting use rates in their role as FCR estimates.

# 2   Methods

## 2.1   Target Population Studied

ADF&G supplied a list of 268 unique rural census-designated communities for inclusion in this analysis.[6] Six of these 268 communities (Chisana, Flat, Port Clarence, Red Dog Mine, Attu Station, and Mertarvik) had zero population (individuals) living in households (according to the 2010 U.S. Census) and were, therefore, excluded, leaving 262 communities that define Alaska's rural community population for this report. Residential units consisting of group quarters (e.g., residences for cannery workers) were not included in the data or analyses presented in this report. Therefore, the estimated mean and 90<sup>th</sup> percentile use rates presented in this report refer to the <u>household</u> population[7] of these 262 rural communities.

---

[6] File provided by ADF&G: *Community Characteristics V1.xlsx*. According to ADF&G the list of communities was prepared from the Alaska Department of Labor Website (http://live.laborstats.alaska.gov/pop/index.cfm) and corresponded to rural census-designated places in Alaska. In addition to the 268 unique rural communities, this file also contained an additional 4 records that represented combinations of some of the 268 unique rural communities. These four records were not needed or used in the analysis.

[7] Excluding the population residing in group quarters.

## 2.2   Data Used

In this report, TMWL estimated the means and 90th percentiles for use rates from data provided by ADF&G (referred to as the "raw use dataset").[8] The raw use dataset includes resource harvest and use information from surveys of 110 ADF&G-defined study communities that contained 118 rural census-designated communities.[9] (Unless explicitly stated as either "ADF&G-defined study communities" or "study communities" the terms "community" or "communities" in this report refer to census-designated communities.) The 110 rural study communities were selected as a non-random sample from the 262 rural communities described above. The raw use dataset lists one or more rows for each combination of a surveyed household and a specific resource (e.g., summer chum salmon, black rockfish, mussels).[10] The survey data are amounts of each resource consumed by an entire household and not amounts consumed by an individual.

ADF&G has informed TMWL that "prior to delivering harvest and use data to TMWL for analysis, ADF&G removed cases of non-human resource use (e.g., fish for dogs); in addition, ADF&G utilizes conversion factors that reflect the parts of resources used for human consumption (e.g., fish headed and gutted)."[11] The data prepared in this way by ADF&G are the starting point for the rates calculated and presented in this report.

ADF&G has described the basis for selecting communities for a survey as follows:

> *The collection of comprehensive harvest data for Alaska communities by ADF&G has been supported primarily by special project funds, reflecting the need for data to address the potential effects of development projects and regulatory issues. Limited State of Alaska General Funds have been used to fill in the gaps of the geographic coverage of community harvest survey data.*[12]

A list of communities for which ADF&G has survey harvest data (as of July 2017) is available online (ADF&G 2017b). From the large collection of surveys carried out by ADF&G, the 118 communities included in the present study are rural communities surveyed between 2009 and 2016. The resource use refers to the study years 2008–2015. ADF&G has published many articles and reports based on their surveys. See the article by Fall (2016) for an example of ADF&G's research.

---

[8] Main files provided: *ALL_Initial_2018_06_07_Part_I.xlsx, ALL_Initial_2018_06_07_Part_II.xlsx, ALL_Initial_2018_06_07_Part_III.xlsx* and *ALL_Initial_2018_06_07_Part_IV.xlsx*.

[9] Among the ADF&G-defined study communities in the harvest data, seven such communities included multiple (two or three) U.S. census-designated communities, as follows (using ADF&G community numbering). Dot Lake Village (community #495) was included in Dot Lake (community #115), Four Mile Road (community #1003) was included in Nenana (community #241), Northway Junction (community #437) and Northway Village (community #438) were included in Northway (community #256), Willow Creek (community #491) was included in Kenny Lake (community #185), Seldovia Village (community #573) was included in Seldovia (community #304), Nabesna (community #235) was included in Slana (community #316), Silver Springs (community #785) was included in Copper Center (community #103). The seven ADF&G-defined study communities, thus, include 15 census-designated communities. The "extra" eight census-designated communities (listed in this footnote) add to the 110 count of study communities to bring the total up to the 118 census communities.

[10] Each resource has a unique numeric code; e.g., the code for summer chum salmon is 111010000.

[11] Quoted from comments by ADF&G staff on 9/6/18.

[12] Email from Marylynne Kostick, ADF&G, 10/23/18.

The subsistence areas of Alaska (commonly referred to in this report as "rural") are defined by statute. ADF&G data collection only occurs in areas defined by the Joint Board of Fisheries and Game as subsistence areas or "rural" (ADEC 2018). The following is an extract from the relevant statute [AS 16.05.258(c)], which shows that the rural (subsistence) areas are defined as the complement of the nonsubsistence areas.[13]

> *(c) ... The boards [Board of Fisheries and the Board of Game], acting jointly, shall identify by regulation the boundaries of nonsubsistence areas. A nonsubsistence area is an area or community where dependence upon subsistence is not a principal characteristic of the economy, culture, and way of life of the area or community. In determining whether dependence upon subsistence is a principal characteristic of the economy, culture, and way of life of an area or community under this subsection, the boards shall jointly consider the relative importance of subsistence in the context of the totality of the following socio-economic characteristics of the area or community:*
> *(1) the social and economic structure;*
> *(2) the stability of the economy;*
> *(3) the extent and the kinds of employment for wages, including full-time, part-time, temporary, and seasonal employment;*
> *(4) the amount and distribution of cash income among those domiciled in the area or community;*
> *(5) the cost and availability of goods and services to those domiciled in the area or community;*
> *(6) the variety of fish and game species used by those domiciled in the area or community;*
> *(7) the seasonal cycle of economic activity;*
> *(8) the percentage of those domiciled in the area or community participating in hunting and fishing activities or using wild fish and game;*
> *(9) the harvest levels of fish and game by those domiciled in the area or community;*
> *(10) the cultural, social, and economic values associated with the taking and use of fish and game;*
> *(11) the geographic locations where those domiciled in the area or community hunt and fish;*
> *(12) the extent of sharing and exchange of fish and game by those domiciled in the area or community;*
> *(13) additional similar factors the boards establish by regulation to be relevant to their determinations under this subsection.*

Only data from surveys in subsistence (rural) areas of Alaska have been used for this report. The following variables from the rural data were used in the analysis for this report and are included in the raw use dataset:[14]

- Household, study community, and region identifiers
- Stratum identifier (relevant in two study communities)[15]
- Resource identifier (a unique numeric code for each resource)
- Weight of resource used (in pounds per year)
- Several variables that indicate whether the resource was harvested and used, given away, or received by the household
- Household size

---

[13] Figure 5 of ADEC 2018 maps the subsistence and nonsubsistence areas in Alaska. The nonsubsistence areas are a small fraction of Alaska's total area.

[14] The dataset also contains other variables not described here. TMWL has only noted variables used in this report.

[15] Two study communities were subdivided by ADF&G for survey purposes. ADF&G referred to each subdivision as a "stratum." The five strata (two in one study community and three in the other) have been handled in data files and in the analysis as if they were separate study communities. The stratum identifier is a code that identifies each stratum.

- Total number of households in the study community

## 2.3 Calculating Use Rates for Households

The ADEC technical workgroup evaluating technical options for developing Alaska HHC selected the species to be included in the FCR used to derive HHC (ADEC 2018). These species included salmon, halibut, herring, non-marine fish, and marine invertebrates. TMWL and ADF&G worked together to calculate use rates per person for these combined resources.[16] ADF&G used data from its study community surveys to calculate the use rate per household for each individual resource. TMWL refers to these use rates per resource, per household as the "raw use data." ADF&G aggregated (added together) the use rates for the individual resources that are included in the target resource group: salmon, non-marine fish, halibut, herring, and marine invertebrates. Starting from the same raw use dataset, TMWL independently calculated the aggregated use rates per household for the combined resource group salmon, non-marine fish, halibut, herring, and marine invertebrates. TMWL and ADF&G jointly verified that their aggregated use rates were identical for every household in the dataset. There are more steps to the data preparation than just described, and the balance of this section and Appendix B provide more detail on the calculation of use rates per household.

ADF&G used methodology developed by Wolfe and Utermohle (2000) to classify households into groups in order to calculate use rates per resource for each household—the "raw use data."[17] TMWL used the Wolfe-Utermohle method of grouping households, because it is tailored to fit the data as collected. (Some comments on the Wolfe-Utermohle method are included in the discussion section of this report.) As part of this methodology, households are classified into one of three groups for a given resource, based on use of that resource:

- Group 1—Selected if a household "harvested and did not give" the resource to other households. The per-person use rate for each household in Group 1 is calculated simply by dividing the household use amount for the resource by the number of household members.
- Group 2—Selected if a household "harvested and gave" the resource to other households or "did not harvest and received" the resource or "did not harvest and did not receive and gave" the resource. (There were households that did not harvest and did not receive but reported giving a specific resource to other households.[18]) In Group 2, the use rate per person is calculated by dividing the grand total use amount of all Group 2 households in the study community by the total number of household members in all households of Group 2.
- Group 3—Selected if a household "did not use" the resource. "Did not use" means they did not harvest or receive or give the resource. They are non-consumers. Group 3 includes only non-consuming households, and the use rate per person for the specific resource is zero.

---

[16] This combined resource group includes multiple individual resources, each with its own unique resource code.
[17] Text in quotes in the bullet points can be found on pages 10–11 of Wolfe and Utermohle 2000.
[18] The non-harvesting, non-receiving households (for a given resource) most likely were able to use and give away some of the resource that was obtained during a year previous to the reference year of the survey inquiry.

Note that a given household may be in a different group when different resources are considered. For instance, a household might be in Group 3 ("did not use") for the Pacific halibut resource and in Group 1 ("harvested and did not give") for summer Coho salmon. Thus, the use rates per resource per household can be quite varied across the resources, even in a single household. In both ADF&G's and TMWL's analyses, the raw use data were aggregated across several specific resources and then the total was divided by the number of household members to yield an estimated average use rate per household member for each surveyed household for the combined resources of salmon, non-marine fish, halibut, herring, and marine invertebrates.[19] The aggregation process is carried out in several detailed steps, and the calculation of the rates per resource (the "raw use data") also requires some steps. See Appendix B for technical details on the calculation of use rates for each household.

## 2.4   Outliers

TMWL carried out an outlier adjustment for each resource in each study community, based on the calculated household-specific use rate of the resource per person. The outlier households for a given resource and study community were then identified using the community-wide grand mean use rate and its standard deviation, along with the ADF&G-suggested upper limit of the daily intake of fish/protein of 340.2 grams/day or, equivalently, 273.94 pounds/year (ADF&G 2017a).[20] For a given resource, an outlier was defined as a use rate that was (1) more than two standard deviations greater than the community-wide mean use rate for the specific resource and (2) that exceeded the threshold of 273.94 pounds/year. See section 4.5 for further discussion of outliers.

The outlier detection is limited to the households assigned to Group 1 for a given resource. Wolfe and Utermohle's Group 2 is assigned a constant per-person use rate within a given study community, and Group 3 has a use rate of zero (non-consumers). Thus, Groups 2 and 3 cannot have outliers. A Group 1 household identified as an outlier was reassigned to Group 2 for the subsequent analysis steps, including recalculation of the Group 2 mean per capita consumption rate for the specific resource after inclusion of the outlier household. A total of 84 outliers[21] were detected and addressed in this way. There were 190,202 household-resource combinations overall—among all surveyed communities. However, only 6,839 household-resource combinations were in Group 1 and subject to screening for outliers. The 84 designated outliers were 1.2% of the 6,839 household-resource combinations in Group 1. A discussion about the impact of the outlier adjustments is included in section 4.5.

## 2.5   Quality Check and Adjustments

To closely replicate ADF&G's implementation of the aggregation and outlier adjustment in its 2017 analysis, TMWL followed methods described in the ADF&G report (ADF&G 2017a, based

---

[19] ADF&G used this aggregation process both for its 2017 report (ADF&G 2017a) and for its cooperative work with TMWL to come up with a final dataset for analysis by TMWL for this report.

[20] The threshold value of 340.2 grams per day is noted as "twice the upper limit of the suggested daily intake of fish/protein" in ADF&G 2017a.

[21] There were a total of 6 outliers for Southeast region households, 8 for Southcentral region households, 9 for Southwest region households, 25 for Western region households, 24 for Arctic region households, and 12 for Interior region households.

on Wolfe and Utermohle 2000) and reviewed ADF&G's corresponding software code.[22] This aggregation process yielded the per capita use rate of salmon, halibut, herring, non-marine fish, halibut, herring, and marine invertebrates (combined)[23] per household, expressed in grams per day. This dataset is referred to as the "aggregated dataset." The aggregated dataset prepared by TMWL was compared to a similar dataset independently prepared by ADF&G. To address differences in the per capita use rates between the TMWL and ADF&G datasets, ADF&G and TMWL staff worked together and followed discrepant cases through the aggregation process and identified changes needed (and then implemented) in TMWL and/or ADF&G codes.[24]

After fixing all codes, the aggregated per capita use rates for salmon, halibut, herring, non-marine fish, and marine invertebrates (combined) were the same for all households in TMWL and ADF&G datasets.

## 2.6   Comparison of Sampled and Non-sampled Communities

In this report, TMWL considered the communities selected by ADF&G for a survey on resource use as a sample drawn from all the rural communities in Alaska.[25] (See section 2.2 of this report, above, for a description of how ADF&G selected communities for its surveys.) TMWL compared the sampled rural communities to the non-sampled rural communities in order to determine to what extent the sampled communities were representative of all rural communities in each region of and in the entire state of Alaska. Thus, community characteristics were compiled and compared between sampled and non-sampled communities both statewide and by region. An Excel presentation of this comparison is included as an attachment to this report (*sampled vs nonsampled communities 09-25-18.xlsx*). A list of the sampled and non-sampled communities can be found in Appendix C, Table C1. The geographical distribution within Alaska of the two groups of communities appears in Figure C1 of Appendix C. The six regions of Alaska are illustrated in Figure C2.

TMWL tabulated and compared mean values, as well as standard deviations and ranges, of continuous-valued characteristics of sampled and non-sampled communities. The statistical significance of the differences between the continuous variables (e.g., population size) in the two groups was determined using the two-sample t-test and the Wilcoxon rank sum test. TMWL also

---

[22] ADF&G used the statistical software SPSS.

[23] These are resources included in Table 4 of ADF&G 2017. Resource codes are described in the file *ADFG_sub_masterCodes.xlsx* provided by ADF&G staff. In that file, resource codes for salmon are 111000000-119000000, marine fish (halibut and herring) (120200000-120310000, 121800000), non-marine fish codes are 124600000-126499000, 129800000-129900000, and marine invertebrates codes are 500200000-509900000.

[24] Changes in code included: (1) Truncate the per capita per year use level at 273.94 pounds in Group 1 in TMWL's code. (2) Use the stated threshold for outlier designation in ADF&G's code. (3) Carry out ADF&G Group 2 aggregations by study community, resource, and strata, and not by study community and resource alone. TMWL changed its code to reflect this. (4) ADF&G's Group 2 included households that did not harvest and did not receive a resource but gave the resource away. These households were previously categorized as Group 3 (non-consumers) by TMWL. TMWL changed its code to place these households in Group 2. The non-harvesting, non-receiving households (for a given resource) most likely were able to use and give away some of the resource that was obtained in year previous to the reference year for which the household survey was conducted.

[25] In common statistical terms, all the rural communities of Alaska taken together can be treated as a population of communities. In this "population" each community is counted once, so there are 262 elements in the population. The population size is 262. From this population, ADF&G drew its sample of 118 communities. The sample size (of communities) is 118.

tabulated and compared categorical characteristics (e.g., categorized population size) of sampled and non-sampled communities. The statistical significance of the differences between the categorical distributions in the two groups was determined using the chi-squared test (and, for comparison within each region, also using Fisher's exact test). The results of the comparison of sampled and non-sampled communities can be found in section 3.1. Other discussion related to this comparison can be found in section 4.4.

## 2.7   Calculation of Use Rates for Alaska and for Regions

TMWL's calculation of the means and the 90th percentiles of resource use rates for regions and for the statewide rural population involves statistical weighting of the household use rates in the surveyed households. These weights adjust the household rates, attempting to mimic the data that would be collected (and subsequently analyzed) if a 100% census of all rural villages (and all their households) were carried out—using the same questionnaires and procedures employed by ADF&G in its survey of the 118 sampled villages.[26] TMWL calculated the statistical weights based on the following three factors.

1.  **The "participation factor."** Communities with a smaller fraction of participating households (which tended to be large communities) had to be up-weighted relative to communities with a larger participating fraction (which tended to be smaller communities, often with an attempt at a full census).

2.  **The size of a household.** There was only one interview per household, but the harvest and use data represented from 1 to 17 persons in the household. The weighting used the household size as a multiplier in the statistical weights.

3.  **A regional adjustment factor.** The third factor adjusted the weights for the uneven representation of sampled villages across the six regions—Table 1 shows the representation of the 118 sampled villages (first numeric column), which can be interpreted as the estimated total population in the communities sampled by ADF&G. The second numeric column is the most recent (2010) census count of the regional populations. The last column (which is the ratio of the first two columns) shows the scale factor that must be used in correcting each region's household rates to mimic the census population count. The Southeast region was highly underrepresented among the 118 villages; note that its adjustment scale factor is the largest among all the regions, with the Southwest region having the second largest adjustment scale factor. Using the U.S. Census totals in Table 1, this adjustment is made by using the method of post-stratification (Levy and Lemeshow 1999).

In summary, the statistical weight assigned to *each household* is the product of the following three adjustment factors, which correspond to the three items described above. The three factors are multiplied together (as noted below) to yield the weight for any given household.

---

[26] The 118 census-designated communities are equivalent to 110 ADF&G-defined study communities. See the section 2.2 for the description of the ADF&G-defined study communities and their relationship to census-designated communities.

1. The participation factor. This is an adjustment for survey coverage of the community. The number of households in the community (enumerated by ADF&G prior to the survey) is divided by the number of households participating in the survey. This factor adjusts for the fact that not all households in a community participated in the survey. The participating households in the community need to be up-weighted to represent all of the households in the community.

*The factor above is then multiplied by:*

2. Household size. The number of household members. This factor gives equal weight to every person in the participating households of a community. This factor reflects the fact that the analysis is intended to estimate means and $90^{th}$ percentiles across _persons_ ("One person, one vote.") and not across _households_ ("One household, one vote.").

*The product of the first two factors is then multiplied by:*

3. A regional adjustment factor. Use the last column of Table 1, choosing the value corresponding to the region of the community. This factor adjusts for the fact that only a certain fraction of the region's rural community population was represented by the communities that ADF&G sampled.

The statistical weight that is the result of these multiplications is specific to and is attached to each household. The weights are used in the computation of the mean and the $90^{th}$ percentiles of the use rates. The weighting is intended to mimic the mean and the $90^{th}$ percentiles of the rates that would be calculated from a full census of all persons from all households in all rural communities.

Absent a census, it is impossible to exactly quantify how successfully the weighting mimicked a full census of all persons in all households in the 262 rural communities. However, if we assume that:

- The participating households in each community are a representative (unbiased) sample of all households in their community, and
- The sampled communities in each region are a representative sample of all communities in their region,

Then, the weighting procedure ensures that the mean and $90^{th}$ percentile estimates of use rates will be a representative (unbiased) characterization of the population (i.e., the estimates will correspond to the hypothetical census values, plus or minus a margin of error due to sampling). Thus, a major strength of the weighted analysis is that it should provide a more accurate representation of use rates across the six regions than an unweighted analysis.

In addition, TMWL employed the statistical weighting in the calculation of margin of error (confidence intervals) and in the comparison of rates between regions. TMWL's survey-

weighted calculations incorporated the complex survey design (one-level cluster sampling[27]) and the additional variability from post-stratification (the regional adjustment factor from Table 1) into the calculation of the 95% confidence intervals for estimated rates. The statistical weights are also used in the calculation of the p-values (calculation of statistical significance) for comparison of regional rates. A more detailed technical description of the weighting can be found in Appendix D.

**Table 1. Calculation of the Regional Adjustment Factor for the Weights**

| Region | The estimated total population in the communities sampled by ADF&G* | Total population** | Regional adjustment factor for the weights |
|---|---|---|---|
| Southeast | 4,038 | 25,328 | 6.272 |
| Southcentral | 6,692 | 7,183 | 1.073 |
| Southwest | 5,103 | 20,438 | 4.005 |
| Western | 13,151 | 24,038 | 1.828 |
| Arctic | 15,353 | 22,979 | 1.497 |
| Interior | 4,798 | 7,501 | 1.563 |

Notes: The regional adjustment factor (in the last column of this table) is the third of the three weighting factors noted in the discussion preceding this table. See the text, above the table, starting with "3. A regional adjustment factor."
*This population figure per region equals the sum of weights before post-stratification (i.e., the participation factor multiplied by the household size).
**Total population living in households in rural communities and excluding population residing in group quarters (U.S. Census 2010).

Appendix D, Table D2 is offered for completeness. The table shows the number of households participating in the surveys used for the rate calculations in this report. The number of household members is also shown in the table, though data were available only per household and not per household member.

## 3  Results

### 3.1  Comparison of Sampled and Non-Sampled Communities

TMWL compared characteristics of sampled and non-sampled communities to determine whether the sampled communities could be considered representative of all communities in a

---

[27] In the one-level cluster sampling design, clusters are first sampled from the population of all clusters (i.e., the set of all rural communities in this study) and then individual observations (i.e., households in this study) are sampled from the collection of all individual observations (i.e., all households in a community) within the sampled communities. As noted, in this study communities represent clusters and households represent individual observations. A consequence of the one-level cluster sampling design is that individual observations in the same cluster may not be statistically independent (i.e., the data from different households in a community may be correlated). The analysis needs to adjust for this within-cluster correlation. If the analysis does not adjust for correlation (e.g., if the data are analyzed as a simple random sample) the actual coverage of the 95% confidence intervals will be less than 95% (and the confidence intervals will overstate the accuracy of the sample estimates).

region. There were large differences among the sampled and non-sampled communities in their proportional distribution across the regions. For example, only 6% of the sampled communities, but 22% of the non-sampled communities, were in the Southeast region, indicating that the Southeast region was under-represented in the sample of 118 rural communities (relative to the population of all rural communities). Southwest and Western region communities were also under-represented, while communities in the Southcentral, Arctic, and Interior regions were over-represented. Overall, the regional composition of the sampled communities differed substantially from the regional composition of all rural communities (the sample was over-represented by communities in the Southcentral, Arctic, and Interior regions and under-represented by communities in the Southeast, Southwest, and Western regions). The difference was very unlikely to be due to chance ($p < 0.001$).

Aside from region, almost all differences in other characteristics that were compared between the sampled and non-sampled communities were not statistically significant in the statewide analysis (the differences were mostly small and possibly consistent with chance). A notable exception was the difference in the presence of a community on a road network—36% of the sampled communities were on a road network, while only 22% of the non-sampled were on a road network (a large and statistically significant difference,[28] $p = 0.008$). However, the "on-road" difference between sampled and non-sampled communities was driven by the different regional composition of the sampled and non-sampled communities. Once the analysis was adjusted for the regional composition (using logistic regression), the presence of communities on a road network was a relatively weak and statistically non-significant predictor of sampled communities ($p = 0.70$).[29] Some statistically significant differences in a few other characteristics were found in individual regions but not statewide. These characteristics can be viewed in the Excel file that accompanies this report (*sampled vs nonsampled communities 09-25-18.xlsx*).

Due to the large number of comparisons between sampled and non-sampled communities across the six regions and for the state as a whole (almost 500 comparisons) and due to the region-specific nature of almost all of the statistically significant differences, the adjustment factors in the last column of Table 1 were the only readily usable results of the comparison of characteristics between sampled and non-sampled communities. Further, accounting for the possibility of true differences in the region-specific analyses would require more complex modeling that was not possible within the scope of this analysis. Although such modeling could potentially improve the weighting and make the analysis of the sampled communities more representative of the regions and the state, the improvement is likely to be only minor, because the differences between sampled and non-sampled communities in the other characteristics were not as large as the differences in the regional composition. A potential drawback of more complex weighting is that it could lead to much wider confidence intervals, indicating more uncertainty in the estimated rates.

---

[28] The large difference can be described by an odds ratio of 2.01. This means that the odds of a community on a road being sampled was about twice the odds of a community not on a road being sampled.

[29] The odds ratio was 1.17. This means that the odds of a community on a road being sampled was 17% more than the odds of a community not on a road being sampled.

## 3.2   Examination of Statistical Weights

The final statistical weights for the analyzed households ranged from 1.11 to 427.5. The diversity of the weights reflects three factors: (1) the differential fraction of communities sampled across the six regions,[30] (2) the varying fraction of households interviewed in each community (out of the total number of households in the community), and (3) the varying household sizes. The impact of the differential representation of communities across the six regions (the first factor mentioned) is shown in Table 1, last column. In that column, the scaling factors used in weighting (that were needed to have the sampled communities represent the total regional population) vary widely across the regions. Relative to the other regions, the populations in the Southeast and Southwest regions were under-represented in the sample and therefore received proportionately higher weights in the analysis.

The final weights were also inversely proportional to the fraction of households (the second factor) interviewed in each community. The fraction ranged from 13.9% to 100% (mean 67.5%) across the 118 analyzed communities. The ADF&G surveys tended to sample a smaller fraction of households (out of the total number of households) in larger communities. Therefore, households from larger communities were under-represented in the sample. To compensate for this under-representation, communities sampled at a lower fraction obtained proportionately larger household weights and final weights.[31]

Finally, household sizes (the third factor mentioned) varied from 1 to 17, which also proportionately contributed to the diversity of the statistical weights. Within a given community, larger households were given more statistical weight than smaller households.

All three factors discussed here entered into the determination of the final weights. The weighting provides more robust and defensible resource use rates for the regions and for the state.

## 3.3   Resource Use Rates for Rural Alaska Communities

Overall and region-specific estimated means and 90th percentiles of the per capita consumption of salmon, halibut, herring, non-marine fish, and marine invertebrates in rural Alaska are shown in Table 2 and in Figures 1–4. These estimates are based on the statistical weighting described in section 2.7 above.

These calculations are based on 118 communities and 6,632 households. Within each sampled community,[32] the number of households participating in the survey represented 13.9%–100% of all households within their community (mean, 67.5%).[33] Among the 6,632 households, a total of 6,118 sample households (92%) used salmon, halibut, herring, non-marine fish, or marine

---

[30] More specifically, this component of the weighting reflects the varying fraction of the population living in the sampled communities in each region.

[31] The relationship of the statistical weights to the community size is illustrated by a strong positive Pearson correlation (r = 0.81) between the total number of households in each community and the household sampling weights adjusted for non-response.

[32] The total noted here counts the strata within a community as separate communities. Specifically, Dillingham had two strata and Kenny Lake had three strata.

[33] By design, less populous communities tended to have larger sampling fractions than more populous communities.

invertebrates, and 514 households (8%) did not use any of them and are therefore referred to as non-consumers.

**Table 2. Per Capita 90th Percentile and Mean Consumption of Salmon, Halibut, Herring, Non-marine Fish, and Marine Invertebrates (Grams Per Day) in Rural Alaska by Region and All Regions**

| | Consumers and Non-consumers | | | | | Consumers-Only | | | | |
| | *P90* | | | *Mean* | | | *P90* | | | *Mean* | |
| *Region* | *rate (95% CI)* | *SE* | | *rate (95% CI)* | *SE* | | *rate (95% CI)* | *SE* | | *rate (95% CI)* | *SE* |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **Southeast** | 318 (140–402) | 67 | | 147 (73–220) | 38 | | 320 (144–410) | 68 | | 152 (79–225) | 37 |
| **Southcentral** | 209 (115–277) | 41 | | 105 (73–136) | 16 | | 217 (117–297) | 46 | | 113 (78–147) | 18 |
| **Southwest** | 287 (100–346) | 63 | | 140 (50–230) | 46 | | 287 (121–350) | 58 | | 145 (56–234) | 45 |
| **Western** | 376 (210–463) | 65 | | 183 (115–251) | 35 | | 379 (216–466) | 64 | | 190 (123–256) | 34 |
| **Arctic** | 275 (113–342) | 58 | | 114 (66–162) | 24 | | 291 (116–350) | 60 | | 125 (76–173) | 25 |
| **Interior** | 235 (211–254) | 11 | | 111 (94–128) | 9 | | 246 (221–266) | 12 | | 127 (110–144) | 9 |
| **All Regions** | 302 (241–350) | 28 | | 141 (111–172) | 16 | | 308 (248–355) | 27 | | 149 (118–180) | 16 |

Notes: See Figures 1–4 for a plot of the P90 and mean estimates.
P90: 90th percentile; CI: confidence interval; SE: standard error



**Figure 1. Per Capita 90th Percentile Use Rate of Salmon, Halibut, Herring, Non-Marine Fish, and Marine Invertebrate (Grams Per Day) Among Consumers and Non-Consumers (Combined) in Rural Alaska Communities by Region (95% Confidence Intervals are Displayed)**

**Figure 2. Per Capita Mean Use Rate of Salmon, Halibut, Herring, Non-Marine Fish, and Marine Invertebrate (Grams Per Day) Among Consumers and Non-Consumers (Combined) in Rural Alaska Communities by Region (95% Confidence Intervals are Displayed)**

**Figure 3. 90th Percentile Use Rate of Salmon, Halibut, Herring, Non-Marine Fish, and Marine Invertebrate (Grams Per Day) Among Consumers Only in Rural Alaska Communities by Region (95% Confidence Intervals are Displayed)**



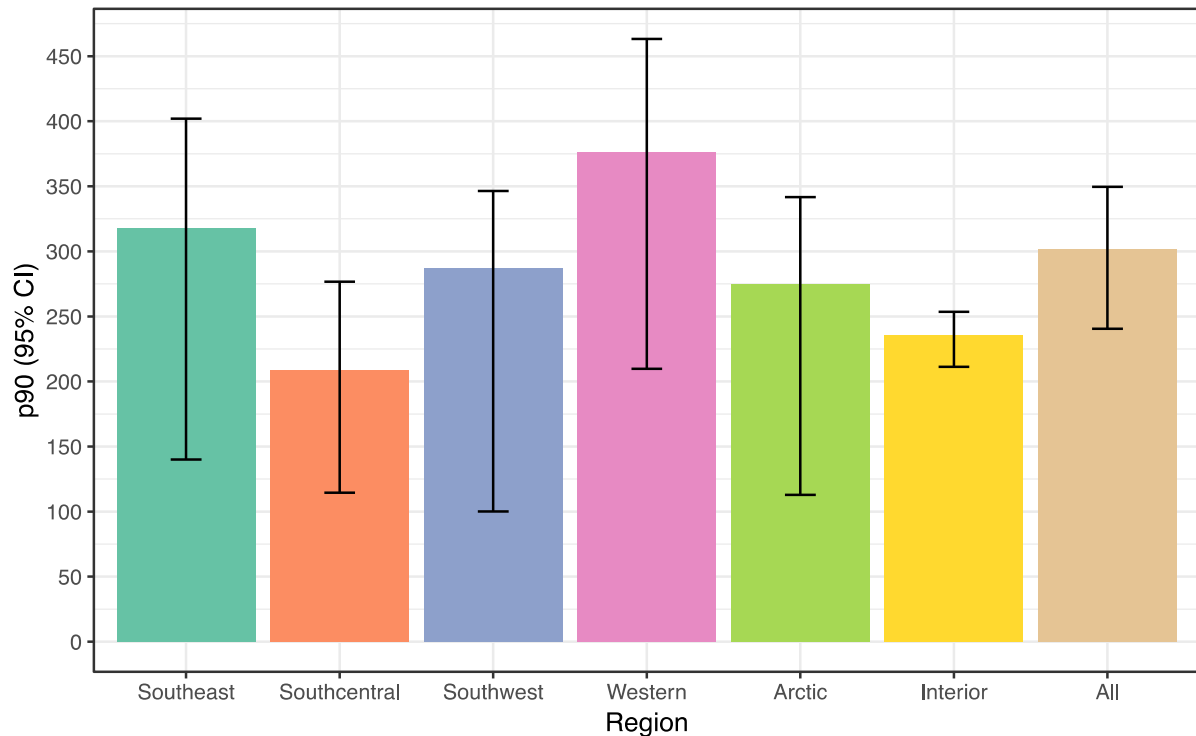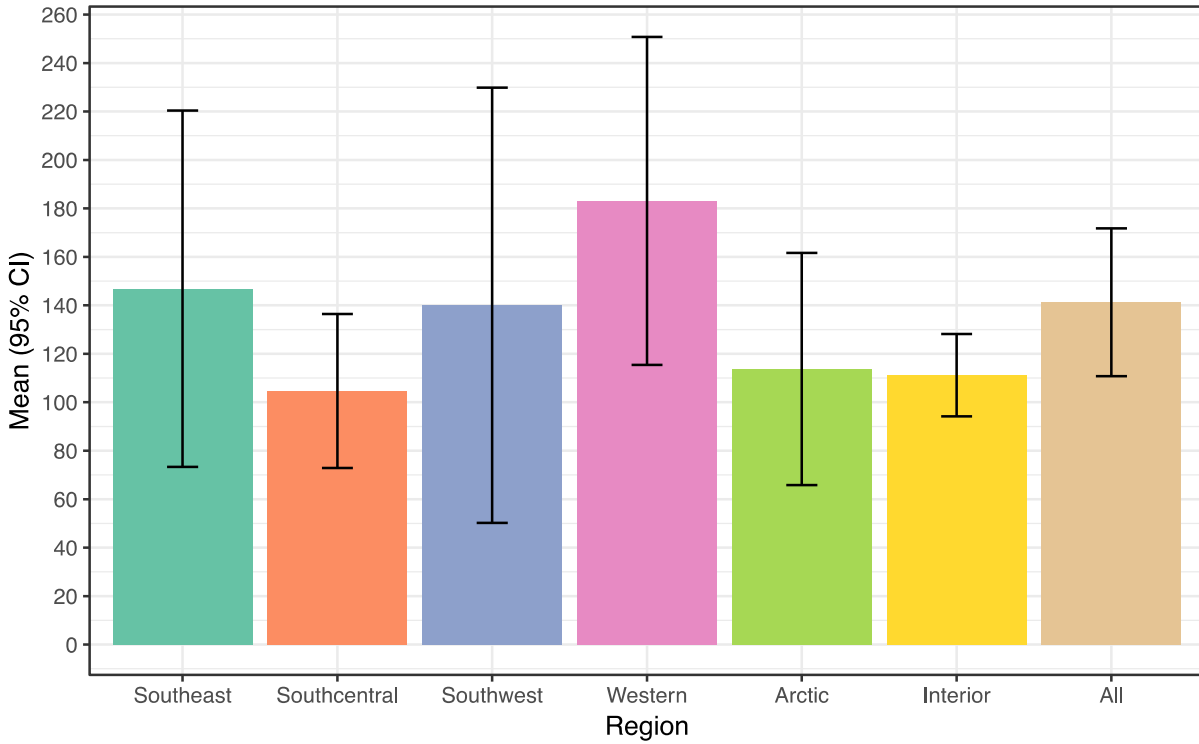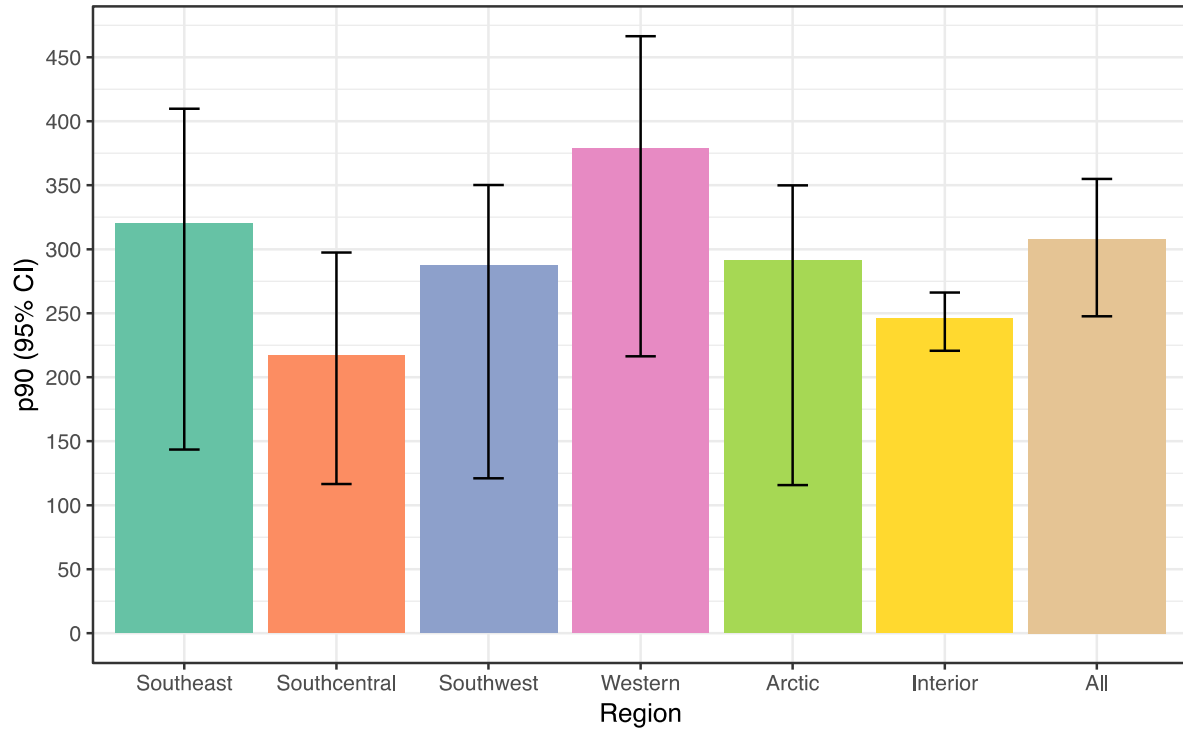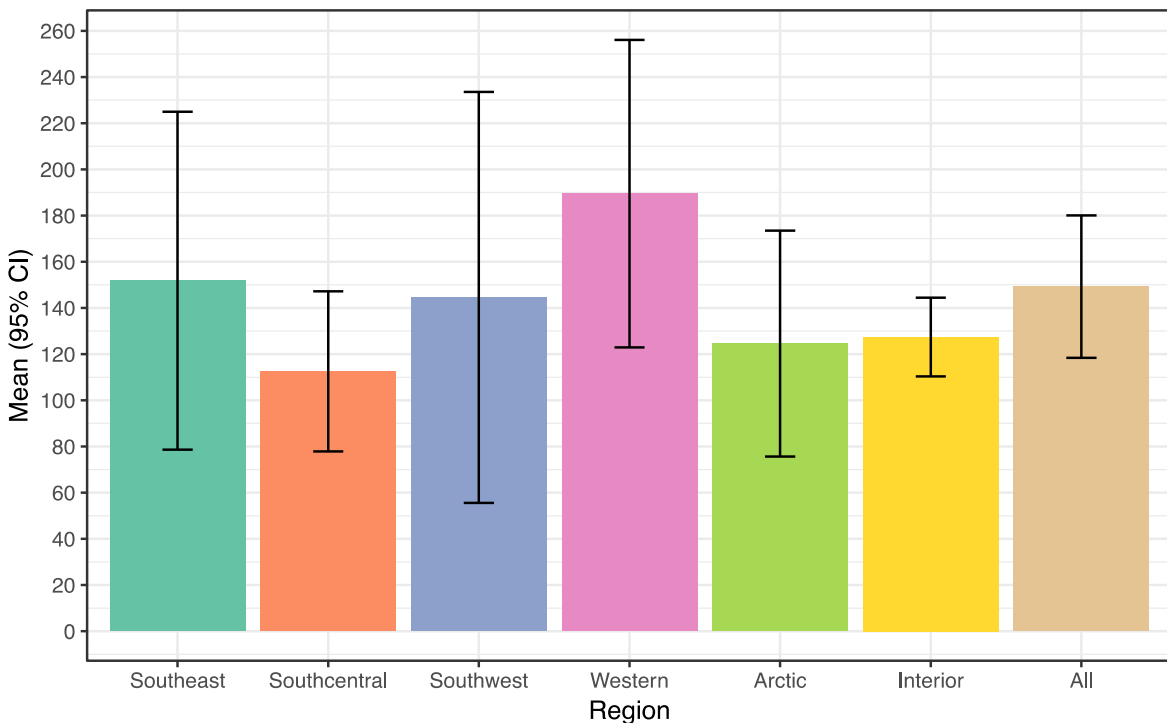**Figure 4. Mean Use Rate of Salmon, Halibut, Herring, Non-Marine Fish, and Marine Invertebrate (Grams Per Day) Among Consumers Only in Rural Alaska Communities by Region (95% Confidence Intervals are Displayed)**

14

Among consumers only, the estimated mean and 90th percentile consumption rates were 308 and 149 grams per day, respectively. The estimated means and 90th percentiles varied substantially among regions, with the Southcentral region and the Western region presenting the largest contrast in summary rates. The smallest 90th percentile was estimated for the Southcentral region (217 grams per day) and the largest for the Western region (379 grams per day, 75% larger). The smallest mean was estimated for the Southcentral region (113 grams per day) and the largest for the Western region (190 grams per day, 68% larger). The differences in the estimated consumer only means and 90th percentiles among the regions were not statistically significant (p = 0.25 for the 90th percentile comparison and p = 0.56 for the mean comparison).

The estimated 90th percentile and mean of the per capita consumption were 302 and 141 grams per day, respectively. The estimated means and 90th percentiles varied substantially among regions, with the greatest contrast occurring again between the Southcentral and Western regions. The smallest 90th percentile was estimated for the Southcentral region (209 grams per day) and the largest for the Western region (376 grams per day, 80% larger). The smallest mean was estimated for the Southcentral region (105 grams per day) and the largest for the Western region (183 grams per day, 74% larger). The differences in the estimated means and 90th percentiles among the regions were not statistically significant (p = 0.19 for the 90th percentile comparison and p = 0.50 for the mean comparison).

The summary rates in Table 2 are quite similar between the consumers only and total (per capita) populations, due to the relatively small fraction (8%) of non-consuming households for this group of resources.

## 3.4   Bias and Sources of Uncertainty

The 90th percentile rates reported here are likely to be underestimates of the true 90th percentile. The assignment of a constant average (mean) rate to every member of every Group 2 household in a community will have reduced some relatively larger within-household and relatively larger within-community individual rates to a lower estimated value. The same reduction in relatively large within-household rates in Group 1 will happen due to the assignment of the same average within-household rate to every member of a given Group 1 household.

In addition, there are likely to be some non-consuming individuals in some households with consumers, and such a household would be designated as a consumer household as long as some individuals do consume the relevant resources. The average use rate for the household (a rate that is used in the calculation of regional and state rates) would be reduced, because the averaging counts all the household members, including the non-consuming (zero consumption) individuals, such as infants. It is likely that in this population the fraction of non-consumers is very small. The impact on use rates of the averaging and the presence of non-consumers in a consuming household cannot be quantified. It could be negligible or larger.

While the regional differences in the 90th percentile rates and in the mean rates were not statistically significant, lack of statistical significance is not an indication of equivalence.[34] Some large regional differences are suggested by the wide variation in estimated means and 90th

---

[34] A common mistake in the scientific literature is the interpretation of lack of statistical significance as an indication of "no difference."

percentiles. The very wide confidence intervals for rates also allow the possibility of large differences among regions, as well as the possibility of smaller regional differences than those estimated in Table 2.

## 4   Discussion

### 4.1   Regional Variation

The estimated use rates presented in this report show substantial regional differences. However, the lack of statistical significance of the regional variation can also be consistent with zero difference in mean and 90th percentile rate among the regions. The numerical evidence leans toward regional variation but does not rule out no variation.

Based on the data that were obtained from ADF&G surveys and TMWL's review of ADF&G printed and Web materials, TMWL noted that a reliable methodology was reported by ADF&G for its field work and data collection. That is a strength supporting the data. Coming up with a distribution of use rates for a population of individuals (by region or in the whole state) is challenging with the survey data as a starting point, because the data do not provide any dietary assessments for individuals.[35] To address that issue, the method of Wolfe and Utermohle (2000) has been used by ADF&G and in this report. That method is a path that leads to rates, but the method calculates a uniformly constant rate for certain groups of individuals. The implication of that methodologic feature is covered in sections 4.2 and 4.6.

### 4.2   Comment on Assumptions Used

All the data on amount of resource used refers to each household collectively, combining all household members. The ADF&G staff have creatively faced this issue by using the methodology developed by Wolfe and Utermohle (2000), which partially lessens the problem. However, a large fraction of individuals (Group 2) in each community are assigned the same (constant) use rate—per resource, regardless, for example, of age or gender. Additionally, every person in any given household is assigned the same household-specific average use rate. That average can vary from household to household in Group 1 households. The effect of these assignments of an average is that true, larger-than-average individual use rates will be biased downward and true, smaller-than-average individual use rates will be biased upward. The distribution of true, individual use rates will be more spread out than the distribution of rates calculated under the Wolfe-Utermohle constant-rate assumptions. And, the reductions from large, true rates to the calculated average rates will reduce the estimated 90th percentile (and other high percentile use rates). Of note, the mean use rate for each community is not biased by

---

[35] An exception would be a dietary assessment for a one-person household.

the "constant" assumption noted above, and, with proper weighting, such as that used in this report, the means for the regions and for the state of Alaska are also unbiased.[36]

## 4.3    A Possible New Survey Feature

ADF&G may wish to consider introducing a feature in their future surveys to check estimates of mean use rate for a region or for the state or for any large-enough collection of communities or households. If a randomly selected individual in each household or in a random sub-sample of households in a community could be interviewed about personal consumption of resources, that data, collectively across communities, could be used as a check on the calculated regional or state (or other) means and percentiles. Such data could be used to assess the downward bias in the 90th percentile (or any high percentile) rates.

## 4.4    Non-random Sampling

The sampled communities are not a random sample of the rural communities of Alaska. The communities were selected and surveyed at different times to address the mission of ADF&G, and the selected communities, pooled together, were not the result of a random selection. ADF&G is to be commended in taking the opportunity to use the data to obtain regional and statewide estimates of consumption. It is very common for data collected for one purpose to be used for another.

As described earlier in this report, TMWL has attempted to correct for the non-random selection of communities by use of appropriate weighting. One prominent component of the weighting addresses the regional distribution of the sampled communities in relation to the total population of rural communities. An extensive comparison of other characteristics between sampled and non-sampled communities left TMWL with only the regional distribution as a characteristic for which to adjust. It is possible, even likely, that other characteristics differ between the sampled and non-sampled communities, but, as noted earlier, characteristics were either specific to one or more regions (but not all regions) or it was not clear that a characteristic had a true (versus observed) difference between sampled and non-sampled communities. Introducing an adjustment based on an uncertain relationship between a characteristic and sampled/non-sampled status could introduce greater uncertainty in the estimates and wider confidence intervals. TMWL had expected that among the several dozen characteristics some would stand out prominently and consistently (across the six regions) as something to use in the analysis. That did not happen. Only the varying fraction of communities sampled per region stood out as an important adjustment factor. It is possible (although not guaranteed) that with additional effort—beyond the scope of this project—the community characteristics could be further used to reduce the potential selection bias. It is also possible that there are unmeasured factors that differ between sampled and non-sampled communities. Such factors, uncontrolled in the weighted statistical analysis, may have biased the estimated mean and 90th percentile use rates away from their true values. However, unmeasured factors might plague <u>any</u> analysis, and there is little basis for this

---

[36] The noted lack of bias of the mean is a lack of bias <u>relative to</u> a population mean that would be obtained if all households in all rural communities were interviewed with the same field methodology as used in the surveys analyzed here. There may be bias that is related to the type and format of questions on harvest—an issue that is not considered here.

concern unless expert opinion backed by valid literature supports the likely existence of such biasing factors in the context of resource use among rural Alaskans.

The non-random selection of communities by ADF&G remains as a non-quantitative source of uncertainty. However, the investigation of community characteristics suggests that there is not an obvious factor, aside from region, that, in a statewide comparison, distinguishes sampled from non-sampled communities. The comparison does suggest that within specific regions, there may be some characteristics that do differ between the sampled and non-sampled communities.

## 4.5   Outliers

In this report, TMWL has adopted the method of outlier designation used by ADF&G.[37] The step-by-step procedure for designating and then handling values as outliers appears in Appendix B. Briefly, a resource-specific use rate (from Group 1 households) that exceeds the community mean for the resource by two or more standard deviations and that exceeds a designated threshold level is assigned the "outlier" label. The method of outlier detection is automatic, whereas statisticians and data analysts often assess outliers with some investigations that are carried out for each outlier and that provide more context and justification for retention or removal of the value. For example, graphics are commonly used (e.g., a histogram or another examination of the distribution of the values). As noted at the end of this section, the handling of outliers in the data analysis appears to have had at most a small impact on the mean and 90th percentile use rates.

There also may be an issue that the ADF&G outlier method includes the requirement of the value exceeding a fixed threshold level that does not vary across resources. Thus, resources that are harvested at a low rate are less likely to be designated as an outlier, since the threshold level lies far beyond a relatively low mean value.

However, the specific outlier method used in this analysis does not appear to have had much effect on the estimated rates compared to the use of outlier values "as is." The automatic method of ADF&G detected a few dozen outliers (among 6,000+ Group 1 use rates that were screened for outliers). Based on a sensitivity analysis, the outlier method has not had a substantial impact on the estimated means and 90th percentile use rates reported (the primary analysis of this report). Compared to the primary analysis (with the outlier adjustment), the "no-outlier" estimated means for the whole State of Alaska differed by 0.1% or less, and the 90th percentiles differed by 1.5% or less. All estimated region-specific means differed by 0.5% or less, and the region-specific 90th percentiles differed by 7% or less from the estimated rates of the primary analysis.[38]

---

[37] ADF&G's process for detecting outliers for the purpose of this report was developed after considering input from the working group members as well as some personal communications with outside professionals (personal communication 12/19/18).

[38] With one exception, the estimated region-specific 90th percentiles differed by 4% or less from the corresponding, outlier-adjusted values. The consumer-only 90th percentile use rate for the Arctic region was 6.5% smaller without outlier adjustment than in the primary (outlier-adjusted) analysis: 272.3 grams per day without outlier adjustment vs. 291.4 grams per day with outlier adjustment.

## 4.6   The Method of Wolfe and Utermohle

The methodology used in this report for calculation of use rates was developed by Wolfe and Utermohle (Wolfe and Utermohle 2000). The method is based on data collected from each of many households. The important data derived from the survey—for this project—is the total amount of each resource used by the members of each household. The total amount used is a household value and not a value attached to a specific person—except in the case of a one-person household.

The household as the unit for data collection has been used by ADF&G for some time, and the sample frame, from which the sample or full census is drawn, is a list of all households in a community. The household design is appropriate when there are questions in the survey that deal with household properties, such as type of facilities, rent versus own, number of rooms, total household income, etc. In addition, a knowledgeable person in the household may be able to address the total household harvest of resources.

A sample of individuals is the ideal for an analysis where the goal is to provide percentiles of resource use for a population of individuals (as in the present study).[39] However, the sampling of households does have operational merits, and the method of Wolfe and Utermohle (2000) does make use of data collected for households. The high percentiles of use (e.g., the 90th percentile use rate) will be biased downward (and low percentiles will be biased upward) due to the assignment of a constant average use rate for each resource for (1) all individuals in each household and (2) all individuals in certain groups of households in a community. Stated equivalently, the distribution of calculated individual use rates based on the Wolfe-Utermohle constant-rate assumption will have upper and lower tails contracted toward the middle of the distribution compared to the distribution of true individual use rates. The extent of bias in percentiles calculated from the Wolfe-Utermohle method cannot be assessed from the data. However, if the true variation in use rates <u>across</u> communities is large relative to the variation in use rates among the households within a community and among the members of a household, then the bias—from the assumption of a constant use rate within each household and a constant use rate among all members of a group of households—may be relatively small.[40]

Given the household-based data for the ADF&G surveys, it is acceptable to calculate use rates by the method of Wolfe and Utermohle. In making use of the ADF&G survey data, the noted bias in percentiles from a household-based survey of resource use does not eliminate the utility of the surveys for estimation of use rates. The bias in percentiles—of unknown magnitude—need not be considered as a barrier to use of the rates and is far from being a fatal survey flaw. Every survey has known and unknown sources of <u>potential</u> bias, and, typically, the magnitude of each type of bias is unknown. Some of the potential biases in the ADF&G surveys are common to almost any survey that elicits responses from individuals and that includes multiple communities. These potential sources of bias include:

- Non-response (non-respondents may differ in resource use from respondents)

---

[39] In the context of the present study, individuals would have to be reached through a sample of households with sub-sampling of one or more individuals in each selected household.

[40] As noted earlier, in a given community all members of Wolfe and Utermohle's Group 2 households are assigned the same average use rate.

- Unmeasured differences between sampled and non-sampled communities—differences that have not been addressed by the statistical weighting
- The unreliability of memory (respondents are reporting on harvesting that occurred during an entire year)

For the collection of ADF&G surveys, the magnitude of the bias in percentiles and from the factors just mentioned cannot be readily estimated. The relative impact of each of these factors is unknown, and none of these factors can easily be singled out as having the most impact or even as having an impact that is not close to zero.

In summary, the ADF&G has collected valuable data on resource use, and the data can be used to estimate resource use rates, such as those presented in this report. The potential sources of bias do not invalidate the survey but can be considered when the rates are used.

## 4.7 Comparison of Alaska's Rural Resource Use Rates with Rates from Other Populations

The resource use rates for Alaska and its regions are generally higher than FCRs of Pacific Northwest and Alaska Native populations and the United States general population (Table 3) as determined by surveys of fish consumption in these populations. The mean and 90th percentile resource use rates for Alaska (and each of its six regions) are larger than most of the corresponding FCRs in Table 3. (See Table 3 and the associated discussion for the exceptions.)

The estimated Alaska statewide resource use rates in Table 3 exceed the estimated FCRs for all of the other populations in the table except the populations covered in the Suquamish and Kodiak surveys. The degree to which Alaska's actual (not estimated) Alaska FCRs exceed the rates for all of the populations (except Suquamish and Kodiak) is likely greater than the tabulation shows, for three reasons, as follows:

- *Different groups of species were combined for calculation of the FCRs for different populations.* Most of the FCRs in the table are for consumption of all freshwater and marine species of fish and shellfish combined.[41] The Alaska resource use rates, on the other hand, are for a more limited set of species designed to support derivation of HHC.[42] Recomputing FCRs for all of the populations covered in Table 3 and using only the more limited group of species covered in the present report would increase the degree by which Alaska resource use rates would exceed the recalculated FCRs of some of the other populations.
- *Some of the FCRs in Table 3 are based on consumption by adults, whereas the Alaska resource use rates in the present report include consumption by both adults and children.* Children would consume less per person, on the average, than adults. If Alaska resource

---

[41] The Nez Perce and Shoshone-Bannock FCRs included rates derived using a more limited set of species designed to support HHC development.

[42] The species included in resource use rates were determined by the ADEC Technical Workgroup supporting HHC development. The workgroup's species selection process was informed, in part, by the EPA FAQ document at: https://www.epa.gov/sites/production/files/2015-12/documents/hh-fish-consumption-faqs.pdf.

use rates were computed for adults only, those rates would exceed some of the other FCRs in Table 3 by a larger margin than now shown.

- *There is a potential downward bias of Alaska's high percentiles of resource use rates*— As discussed in sections 3.4 and 4.2 of this report, there is a bias due to the ADF&G survey's assessment of resource use rates at the household level rather than at the individual level. If Alaska's resource use data were available at the individual level, calculated 90th and other high percentiles of resource use rates would very likely exceed some of the corresponding Table 3 FCRs by a greater extent than what is shown.

As noted in the foreword of this report, resource use rates from the ADF&G surveys were used to calculate FCRs for Alaska HHC development because there was a lack of other fish consumption survey data for the State as a whole. Two surveys of fish consumption have been conducted by Alaska Native tribes: one survey of villages near Cook Inlet (Seldovia Village Tribe 2013), the other survey of villages on Kodiak Island by the Sun'aq Tribe of Kodiak (Lance et al. 2019). See Table 3 and associated notes for additional information on the final draft Kodiak report.

The mean and 90th percentile consumption rates for Alaska (all regions combined, Table 3) are larger than the corresponding rates for all other populations in Table 3 with the exception of the Suquamish Tribe and the Kodiak Island Alaska Native Villages. Among the Alaska populations in Table 3, the average total seafood FCR for the Cook Inlet villages is less than the Alaska statewide average resource use rates (all regions combined).[43] The average and 90th percentile total seafood FCRs for the Kodiak villages exceed the corresponding Alaska resource use rates and are the largest average and 90th percentile rates in Table 3.

As was the case for many of the non-Alaska FCR surveys, the Cook Inlet (Seldovia Village Tribe 2013) and Kodiak (Lance et al. 2019) survey reports did not include FCRs for the subset of species used for HHC derivation. Given the absence of the information on a species group relevant for HHC derivation, the total seafood consumption rates (all seafood species combined) from the Cook Inlet and Kodiak studies were the only available relevant rates for comparison to average Alaska statewide resource use rates.[44] If FCRs from these studies, as well as results from other Region 10 Tribes (e.g., the Suquamish Tribe), were recalculated using the smaller number of species used for calculation of Alaska's resource use rates in the present study, the resulting, limited-species FCRs would be lower. Such a recalculation would result in an increase in the margin by which the average Alaska statewide resource use rates (from the present report) exceed the recalculated average Cook Inlet FCR. Similarly, that recalculation would decrease the

---

[43] The Cook Inlet report did not include a 90th percentile total seafood FCR, and so it was not possible to compare a Cook Inlet 90th percentile rate to the 90th percentile rates for the other populations noted in Table 3.

[44] Fish and non-fish species evaluated for consumption in the Cook Inlet Alaska Native survey may be found at: http://svt.org/wp-content/uploads/2016/03/assessment-of-cook-inlet-tribes-subsistence-consumption-final.pdf. Fish species are tabulated in section 7.4.2.4. Shellfish, bird, and marine mammal species are tabulated in section 7.4.3.1. Fish and non-fish species evaluated for consumption in the Kodiak survey are listed in Appendix III of Lance 2019, Table A4. Specifically, the noted Table A4 lists all of the species evaluated in the FFQ portion of the Kodiak survey. Fish and non-fish species evaluated for consumption in the Kodiak survey are itemized in Table 9 of the draft report (Lance et al. 2019).

margin by which the average Suquamish and draft Kodiak average FCRs exceed the Alaska statewide resource use rates.

**Table 3. Total FCRs (grams/day) of Adults for Rural Alaska Communities, Pacific Northwest Tribes (with Consumption Rates Available) and the U.S. General Population. Consumers Only.**

| Population, age; species consumed | No. of consumer respondents | Mean | 90th percentile |
|---|---|---|---|
| **Alaska's six regions, rural communities, all ages; Salmon, Halibut, Herring, Non-marine Fish, and Marine Invertebrates** | | | |
| All regions combined | 6,632+ | 149 | 308 |
| Southeast | 499+ | 152 | 217 |
| Southcentral | 1,218+ | 113 | 287 |
| Southwest | 645+ | 145 | 379 |
| Western | 1,550+ | 190 | 291 |
| Arctic | 1,663+ | 125 | 246 |
| Interior | 1,057+ | 127 | 308 |
| **Other Native Populations, age ≥ 18 unless noted; all finfish and shellfish unless noted** | | | |
| Nez Perce Tribe[45]; near coastal/estuarine/freshwater/anadromous finfish and shellfish | | | |
| Food frequency questionnaire# | 446 | 104 | 231.4 |
| National Cancer Institute (NCI) method# | 446 | 66.5 | 159.4 |
| Shoshone-Bannock Tribes[46]; near coastal/estuarine/freshwater/anadromous finfish and shellfish | | | |
| Food frequency questionnaire# | 225 | 111 | 266 |
| NCI method# | 225 | 18.6 | 48.9 |
| Tulalip Tribes[47] | 73 | 82.2 | 193.4 |
| Squaxin Island Tribe[48] | 117 | 83.7 | 205.8 |
| Suquamish Tribe[49], age ≥ 16 | 92 | 213.9 | 489.0 |
| Columbia River Tribes[50] | 464 | 63.2 | 130.0 |
| Alaska Native Villages, Cook Inlet, Alaska[51] | 75 | 107.7 | NA |

---

[45] See Polissar et al. 2015a and 2015b, Tables 14 and 28 of each report.

[46] See Polissar et al. 2015a and 2015b, Tables 14 and 28 of each report.

[47] See Polissar et al. 2014.

[48] See Polissar et al. 2014.

[49] See Suquamish Tribe 2000 (Table T-3) and Liao 2002; these rates were converted from g/kg/day to g/day by multiplying by the mean body weight of 79.0 kg, found in Table T-2 of Suquamish 2000.

[50] See CRITFC 1994; The noted "90th" percentile value (130 g/day) is actually the 91.6th percentile from Table 10 of the CRITFC report. The 90th percentile value from linear interpolation using Table 10 is 113.1 g/day.

[51] The average (mean) rate for the Cook Inlet villages (combined) is from Seldovia 2013, Appendix B, value for consumers only, "Total Seafood (snails not included for Seldovia)." The 90th percentile is not shown; the 95th percentile is 267.6 g/day.

| Population, age; species consumed | No. of consumer respondents | Mean | 90th percentile |
|---|---|---|---|
| Alaska Native Villages, Kodiak Island, Alaska[52] | 326 | 232.8[53] | 528.3[53] |
| **USA, general population** | | | |
| NCI method[54], age ≥ 21 years; all finfish and shellfish | 16,363* | 23.8 | 52.8 |

Notes: *The count for the USA includes both consumers and non-consumers, though the fraction of non-consumers is likely to be small. For a discussion of this point, see a footnote to Table 28 of Polissar et al. 2015a or 2015b.

+The Alaska counts in this column refer to the responding households which provided useable data for the rate calculations. The Alaska surveys provided data per species for the entire household and not for individual household members.

#FCRs are based on a subgroup of consumed species.

Except for the rates calculated by the NCI method for the USA and the Nez Perce Tribe and Shoshone-Bannock Tribes, all of the rates in this table are based on some variation of a food frequency questionnaire. The harvest questionnaire of the Alaska surveys also falls in this questionnaire "family," by asking respondents about their harvest over a one-year recall period.

# 5    Conclusions

Based on survey data from 118 rural communities in Alaska, appropriate statistical weighting has provided summary use rates (mean and 90th percentile) for consumers only and for consumers and non-consumers combined (per capita rates) for the rural population of Alaska, consisting of 262 rural communities. The rates represent use of salmon, halibut, herring, non-marine fish, and marine invertebrates (combined). Statewide mean and 90th percentile use rates are very similar between consumers and the total rural population (consumers and non-consumers combined). The statewide consumer-only mean rate for rural communities is 149 grams/day (95% confidence interval, 118–180 grams/day), and the 90th percentile rate is 308 grams/day (95% confidence interval, 248–355 grams/day). The mean and the 90th percentile rates vary substantially among Alaska's six regions, and the rates have a wide margin of error. The differences in summary rates among regions are not statistically significant, but the differences are large enough that they may be judged as practically significant.

A comparison of several dozen demographic and other characteristics between the sampled and non-sampled rural communities showed a large and statistically significant difference in the

---

[52] The rates for the Kodiak villages are from Lance et al. 2019, Table 30, "All Seafood." The rates are based on consumers only.

[53] *This footnote has been authored by EPA.* EPA Region 10 provided grant funding as well as technical support for development and implementation of the Sun'aq survey and also had access to preliminary data. The 2019 draft final version of the report contains well reviewed total FCRs based on food frequency questionnaire methodology used for almost all other referenced fish consumption surveys. ADEC did not have similar access to the Sun'aq report for review and comment at the time the present technical report was drafted. ***Consequently, discussion of how Sun'aq FCRs compare to Alaska resource use rates must be regarded as preliminary until ADEC has had adequate time to review the draft final Sun'aq report.***

[54] The USA NCI method rates are from USEPA 2014, Appendix E, Table E-1. The National Cancer Institute has developed a statistical modeling approach to utilize short term dietary recall data to estimate the usual long-term consumption rate of specific foods to support research relating diet and disease.

proportion of communities sampled among the six regions. After taking account of this regional variation, no other characteristic (demographic or otherwise) differed between sampled and non-sampled communities in a way that clearly justified using the characteristic in the calculation of use rates. The regional adjustment to the statistical weighting is expected to address—wholly or in part—the effect of non-random selection of the 110 study communities and provide support for use of the calculated summary rates in water quality regulation.

ADF&G documents describe a strong and valid methodology for their survey field work and data collection. The method of Wolfe and Utermohle (2000) used in the analysis accommodates the data collected per household but may have introduced a bias in percentiles. The potential sources of bias in the surveys—some are common to any survey—cannot be quantified and do not prevent use of the rates to represent the resource consumption rates of the rural population of Alaska.  With few exceptions, Alaska fisheries resource use rates are somewhat greater than FCRs for other Native American populations within the Pacific Northwest and Alaska (Table 3). This analysis of fisheries resource use rates based on ADF&G community harvest surveys represents a useful gain in knowledge about resource use in rural Alaska.

# 6   References

ADF&G. 2017a. *Preliminary Regional Analysis of Fish Consumption Rate Estimates for Rural Alaska Populations*. Prepared by Alaska Department of Fish & Game, Division of Subsistence, for the Human Health Criteria Technical Workgroup discussion, March 2017.

ADF&G. 2017b. *Overview of Availability of Comprehensive Harvest Survey Data for Alaska Communities*. Alaska Department of Fish and Game Division of Subsistence, July 2017. Available online at: http://www.adfg.alaska.gov/static-sub/CSIS/PDFs/Overview%20of%20Availability%20of%20Comprehensive%20Harvest%20Survey%20Data%20for%20Alaska%20Communities%202017.pdf.

ADEC. 2018. *Evaluation of Key Elements and Options for Development of Human Health Criteria. Technical Workgroup Report*. November 13, 2018. FINAL DRAFT. Prepared by Alaska Department of Environmental Conservation, Division of Water. Available online at: https://dec.alaska.gov/water/water-quality/human-health-criteria/.

Binder, D.A. 1991. Use of estimating functions for interval estimation from complex surveys. *Proceedings of the ASA Survey Research Methods Section* 34–42.

CRITFC. 1994. *A Fish Consumption Survey of the Umatilla, Nez Perce, Yakama, and Warm Springs Tribes of the Columbia River Basin*. Technical Report 94-3. Columbia River Inter-Tribal Fish Commission.

Fall. 2016. Regional patterns of fish and wildlife harvests in contemporary Alaska. *Arctic* 69(1): 47–64.

Lance, T.A., K. Brown, K. Drabek, K. Krueger, and S. Hales. 2019. *Kodiak Tribes Seafood Consumption Assessment: Draft Final Report,* Sun'aq Tribe of Kodiak, Kodiak, AK. Available online at: http://sunaq.org/wp-content/uploads/2016/09/Kodiak-Tribes-Seafood-Consumption-Assessment-DRAFT-Final-Report-26Feb19-FINAL.pdf

Levy, P.S., and Lemeshow S. 1999. *Sampling of Populations, third edition*. John Wiley & Sons, New York, NY.

Liao, S. 2002. Excel Spreadsheets of Percentiles of Consumer-Only Rates (g/kg-day) for the Suquamish Tribe—Various Species Groups.

Lumley, T. 2004. Analysis of complex survey samples. *Journal of Statistical Software* 9(1):1–19.

Lumley, T. 2019. Survey: analysis of complex survey samples. R package version 3.33.

Polissar, N.L., M. Neradilek, D.S. Hippe, A.Y. Aravkin, P. Danaher, and J. Kalat. 2014. *Statistical Analysis of National and Washington State Fish Consumption Data*. The Mountain-Whisper-Light Statistics.

Polissar, N.L., A. Salisbury, C. Ridolfi, K. Callahan, M. Neradilek, D.S. Hippe, and W.H. Beckley. 2015a. *A Fish Consumption Survey of the Nez Perce Tribe*. The Mountain-Whisper-Light Statistics, Pacific Market Research, Ridolfi, Inc. Available online at: https://www.epa.gov/sites/production/files/2017-01/documents/fish-consumption-survey-nez-perce-dec2016.pdf.

Polissar, N.L., A. Salisbury, C. Ridolfi, K. Callahan, M. Neradilek, D.S. Hippe, and W.H. Beckley. 2015b. *A Fish Consumption Survey of the Shoshone-Bannock Tribes*. The Mountain-Whisper-Light Statistics, Pacific Market Research, Ridolfi, Inc. Available online at: https://www.epa.gov/sites/production/files/2017-01/documents/fish-consumption-survey-shoshone-bannock-dec2016.pdf.

R Core Team. 2017. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Available online at: https://www.R-project.org/.

Seldovia Village Tribe. 2013. *Assessment of Cook Inlet Tribes Subsistence Consumption*. Prepared by: Seldovia Village Tribe Environmental Department.

Suquamish Tribe. 2000. *Fish Consumption Survey of The Suquamish Indian Tribe of The Port Madison Indian Reservation, Puget Sound Region*. Suquamish, WA.

USEPA. 2014. *Estimated Fish Consumption Rates for the U.S. Population and Selected Subpopulations (NHANES 2003–2010)*. EPA-820-R-14-002. U.S. Environmental Protection Agency, Washington, DC. Available online at: https://www.epa.gov/sites/production/files/2015-01/documents/fish-consumption-rates-2014.pdf.

Wolfe, R.J., and C.J. Utermohle. 2000. *Wild Food Consumption Rate Estimates for Rural Alaska Population*s. Technical Paper No. 261. Prepared for the Alaska Department of Environmental Conservation, Contaminated Sites Program by the Alaska Department of Fish and Game Division of Subsistence, Juneau, AK.

# Appendix A. Resources (Species) Included in Use Rates

The following table itemizes the resources (and numeric resource codes) which fall under the rubric "salmon, halibut, herring, non-marine fish, and marine invertebrates." Not all of the species itemized in the table actually occurred as consumed resources in the data collected from the 118 sampled villages.

**Table A1. Specific Resources and Their Numeric Resource Codes**

*An asterisk indicates a resource that was reported as used by at least one household, based on the harvest data from the 118 sampled villages. Key: S = salmon; N = non-marine fish; I = marine invertebrates; MF = marine fish (halibut and herring)

| resource number | resource name | key | resource number | resource name | key | resource number | resource name | key | resource number | resource name | key | resource number | resource name | key |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 111000000* | Chum salmon | S | 125002000 | Arctic char | N | 126404000* | Broad whitefish | N | 500899000* | Unknown cockles | I | 502400000* | Oyster | I |
| 111010000 | Summer chum salmon | S | 125004000* | Brook trout | N | 126406000 | Cisco | N | 501000000 | Crabs | I | 502402000 | Rock oyster | I |
| 111020000 | Fall chum salmon | S | 125006000* | Dolly Varden | N | 126406020* | Arctic cisco | N | 501002000 | Box crab | I | 502499000 | Unknown oyster | I |
| 111090000 | Unknown chum salmon | S | 125006010 | Dolly Varden (freshwater) | N | 126406040* | Bering cisco | N | 501004000* | Dungeness crab | I | 502600000 | Scallops | I |
| 112000000* | Coho salmon | S | 125006020 | Dolly Varden (saltwater) | N | 126406060* | Least cisco | N | 501006000* | Hair crab | I | 502602000* | Weathervane scallops | I |
| 112020000 | Coho salmon (fingerling) | S | 125006030 | Dolly Varden (Togiak trout) | N | 126406990* | Unknown cisco | N | 501008000* | King crab | I | 502604000* | Rock scallops | I |
| 113000000* | Chinook salmon | S | 125006990 | Dolly Varden (unknown) | N | 126408000* | Humpback whitefish | N | 501008020* | Blue king crab | I | 502699000* | Unknown scallops | I |
| 114000000* | Pink salmon | S | 125008000 | Dolly Varden (fingerling) | N | 126410000* | Lake whitefish | N | 501008040* | Brown king crab | I | 502800000* | Sea anemone | I |
| 115000000* | Sockeye salmon | S | 125010000* | Lake trout | N | 126412000* | Round whitefish | N | 501008080* | Red king crab | I | 503000000* | Sea cucumber | I |
| 116000000* | Landlocked salmon | S | 125099000* | Unknown char | N | 126499000* | Unknown whitefishes | N | 501008100 | Red king crab eggs | I | 503002000 | Football sea cucumber | I |
| 117000000 | Spawnouts | S | 125200000* | Arctic grayling | N | 129800000 | Any freshwater non-salmon | N | 501008300* | Hanasaki crab | I | 503004000 | Yein sea cucumber | I |
| 117010000 | Spawning chum salmon | S | 125400000 | Northern pike | N | 129900000* | Unknown non-salmon fish | N | 501008990 | Unknown king crab | I | 503006000 | Red sea cucumber | I |
| 117020000* | Spawning Coho salmon | S | 125402000 | Northern pike (large) | N | 500200000* | Abalone | I | 501010000 | Korean horse hair crab | I | 503099000* | Unknown sea cucumber | I |
| 117040000* | Spawning pink salmon | S | 125404000* | Northern pike (small, pickle) | N | 500400000* | Chitons (bidarkis, gumboots) | I | 501012000 | Tanner crab | I | 503200000 | Sea urchin | I |
| 117050000* | Spawning sockeye salmon | S | 125499000 | Unknown pike | N | 500404000* | Red (large) chitons | I | 501012020* | Tanner crab, bairdi | I | 503202000* | Green sea urchin | I |
| 117090000* | Unknown salmon spawnouts | S | 125500000* | Northern pike | N | 500408000* | Black (small) chitons | I | 501012040* | Tanner crab, opillio | I | 503204000* | Red sea urchin | I |
| 118000000 | Salmon roe | S | 125600000* | Sheefish | N | 500499000* | Unknown chitons | I | 501012990* | Unknown tanner crab | I | 503206000* | Purple sea urchin | I |
| 119000000* | Unknown salmon | S | 125800000* | Sturgeon | N | 500600000 | Clams | I | 501099000* | Unknown crab | I | 503208000 | Black sea urchin | I |
| 120200000* | Pacific herring | MF | 125802000 | Green sturgeon | N | 500602000* | Butter clams | I | 501200000* | Geoducks | I | 503299000* | Unknown sea urchin | I |
| 120300000 | Pacific herring roe | MF | 125804000 | White sturgeon | N | 500604000* | Freshwater clams | I | 501400000 | Giant scale worm | I | 503400000* | Shrimp | I |
| 120302000 | Pacific herring roe/unspecified | MF | 125899000 | Unknown sturgeon | N | 500606000* | Horse clams | I | 501600000 | Jingles | I | 503600000* | Snails | I |
| 120304000 | Pacific herring sac roe | MF | 126000000* | Longnose sucker | N | 500608000* | Pacific littleneck clams | I | 501602000* | Rock jingles | I | 503700000 | Starfish | I |
| 120306000 | Pacific herring spawn on kelp | MF | 126200000 | Trout | N | 500610000* | Pinkneck clams | I | 501699000 | Unknown jingles | I | 503800000* | Squid | I |
| 120308000 | Pacific herring roe on hair seaweed | MF | 126202000* | Cutthroat trout | N | 500612000* | Razor clams | I | 501800000* | Limpets | I | 504000000* | Whelk | I |
| 120310000 | Pacific herring roe on hemlock branches | MF | 126204000* | Rainbow trout | N | 500614000* | Softshell clams | I | 502000000 | Mussels | I | 509900000* | Unknown marine invertebrates | I |
| 121800000* | Pacific halibut | MF | 126206000* | Steelhead | N | 500699000* | Unknown clams | I | 502002000* | Blue mussels | I | | | |
| 124600000 | Alaska blackfish | N | 126299000* | Unknown trout | N | 500800000 | Cockles | I | 502004000* | Brown mussels | I | | | |
| 24800000* | Burbot | N | 126400000 | Whitefishes | N | 500802000* | Basket cockles | I | 502099000* | Unknown mussels | I | | | |
| 125000000* | Char | N | 126402000 | Alaska whitefish | N | 500804000* | Heart cockles | I | 502200000* | Octopus | I | | | |

# Appendix B. Aggregation of Resource Use Rates

The steps below present a succinct description of the aggregation process used to calculate the use rate for salmon, halibut, herring, non-marine fish, and marine invertebrates. Note: the term "harvest" (below) refers to "use" when it applies to rates. The "harvest" labeling has not been changed here to keep it consistent with codes used in the R software.

**Step 1. Combine datasets created by ADF&G.** Combine files:

- ALL_Initial_2018_06_07_Part_I.xlsx
- ALL_Initial_2018_06_07_Part_II.xlsx
- ALL_Initial_2018_06_07_Part_III.xlsx
- ALL_Initial_2018_06_07_Part_IV.xlsx

**Step 2. Exclude communities not in the population (non-rural communities).** Data for Denali Park, Ferry, Healy, Trapper Creek, Talkeetna, and Mentasta Pass were excluded.

**Step 3. Combine records into 1 record per household and resource.**

- Harvest weight = sum of harvest weights
- Household size = maximum household size
- Use maxima for indicators
  - Giveaway = max(giveaway)
  - Received = max(received)
  - Harvested = max(harvested)

**Step 4. Define preliminary use group (before the outlier check).** This step assigns a preliminary designation (per household) of the Wolfe and Utermohle groups 1, 2, and 3 (Wolfe and Utermohle 2000).

- PrelimUserGroup = 1 if harvested and did not give away
- PrelimUserGroup = 2 if gave away or received
- PrelimUserGroup = 3 if did not harvest and did not receive and did not give away

**Step 5. Define preliminary use level (before the outlier check).** This step assigns a preliminary use rate per person.

- PrelimUseLevel = harvestLbs_MR/hhSize if PrelimUserGroup = 1
- PrelimUseLevel = sum(harvestLbs_MR)/sum(hhSize) if PrelimUserGroup = 2 (sums calculated by community, stratum and resource)
- PrelimUseLevel = 0 if PrelimUserGroup = 3

**Step 6. Outlier check and adjustment.** First calculate community mean and standard deviation harvest level (by community, strata and resource, including all households in the community). Mark a record an outlier if:

- PrelimUserGroup = 1 AND
- PrelimUseLevel >= community mean + 2 x (community SD) AND

- PrelimUseLevel >= 273.94 lbs.[55]

If the value is marked as an outlier, designate the household as giving the resource away (this will change the user group designation from Group 1 to Group 2 in the next step).

**Step 7. Define the final use group.**
- UserGroup = 1 if harvested and did not give away
- UserGroup = 2 if gave away or received
- UserGroup = 3 if did not harvest and did not receive and did not give away

NA (missing) value was created in UserGroup in the following (relatively very rare) scenarios (1 = yes, 0 = no):

| | harvested | received | giveaway |
|---|---|---|---|
| 1: | 1 | NA | NA |
| 2: | 0 | NA | NA |
| 3: | 1 | 0 | NA |

These records will subsequently have NA (missing) values for the use level (UseLevel), which are effectively counted as zero pounds when the use levels are aggregated across resources within a household.

**Step 8. Define final use level.**
- UseLevel=harvestLbs_MR/hhSize if UserGroup = 1
- UseLevel=sum(harvestLbs_MR)/sum(hhSize) if UserGroup = 2 (sums calculated by community, stratum and resource). If UseLevel > 273.94 lbs. set to 273.94 lbs.
- UseLevel =0 if UserGroup = 3

**Step 9. Truncate use level at 273.94 lbs.**
- If UseLevel > 273.94 lbs. set it to 273.94 lbs.

**Step 10. Create an indicator of resources indicating resources included in Table 4 of the ADF&G report** (salmon, halibut, herring, non-marine fish and marine invertebrates)**:**
- Salmon codes are 111000000-119000000.
- Marine fish codes (halibut and herring) are 120200000-120310000 and 121800000.
- Non-marine fish codes are 124600000-126499000 and 129800000-129900000.
- Marine invertebrates codes are 500200000-509900000.

**Step 11. Aggregate all records for each household.** Aggregate use level per household = sum of use levels that are ADF&G report Table 4 resources (or 0 if none)

---

[55] The raw use data were supplied to TMWL as pounds per year for each resource.

# Appendix C. Sampled and Non-sampled Communities

Table C1 lists the 118 sampled and the 144 non-sampled communities. The resource use rates presented in this report for salmon, halibut, herring, non-marine fish, and marine invertebrates are based on the ADF&G survey data for the 118 sampled communities.

Figure C1 shows the location of the sampled and non-sampled communities within the state of Alaska. The map shows the lower representation of sampled communities (relative to the non-sampled communities) in the Southeast and Southwest areas of the state. The relative preponderance of sampled communities (compared to non-sampled communities) in the northern part of the state is also evident. Figure C2 shows Alaska and its six regions.

**Table C1. List of Sampled and Non-sampled Rural Communities**

### Sampled

| | | | | | |
|---|---|---|---|---|---|
| Akiak | Coldfoot | Gulkana | Mountain Village | Pilot Point | Stevens Village |
| Akutan | Copper Center | Haines | Nabesna | Pilot Station | Susitna |
| Alatna | Cordova | Hoonah | Nanwalek | Point Hope | Takotna |
| Ambler | Deering | Hughes | Napakiak | Point Lay | Tatitlek |
| Anderson | Dillingham | Hydaburg | Napaskiak | Port Graham | Tazlina |
| Angoon | Diomede | Kenny Lake | Nelchina | Quinhagak | Togiak |
| Anvik | Dot Lake | Kiana | Nenana | Rampart | Tok |
| Barrow | Dot Lake Village | Klukwan | New Stuyahok | Ruby city | Tolsona |
| Beaver | Dry Creek | Kobuk | Nikolai | Russian Mission | Tonsina |
| Bethel | Eek | Koliganek | Noatak | Scammon Bay | Tuluksak |
| Bettles | Egegik | Kotzebue | Noorvik | Selawik | Tuntutuliak |
| Buckland | Ekwok | Kwethluk | Northway | Seldovia | Tyonek |
| Cantwell | Emmonak | Lake Louise | Northway Junction | Seldovia Village | Ugashik |
| Chase | Evansville | Manley Hot Springs | Northway Village | Shageluk | Wainwright |
| Chenega | Four Mile Road | Marshall | Nuiqsut | Shishmaref | Whale Pass |
| Chignik | Gakona | McCarthy | Nulato | Shungnak | Willow Creek |
| Chignik Lagoon | Galena | McGrath | Oscarville | Silver Springs | Wiseman |
| Chignik Lake | Glennallen | Mendeltna | Paxson | Skwentna | Yakutat |
| Chistochina | Golovin | Mentasta Lake | Perryville | Slana | |
| Chitina | Grayling | Minto | | Stebbins | |
| Clarks Point | | | | | |

## Not Sampled

| | | | | | |
|---|---|---|---|---|---|
| Adak | Cold Bay | Hyder | Levelock | Pelican | Tanacross |
| Akhiok | Covenant Life | Igiugig | Lime Village | Petersburg | Tanana |
| Akiachak | Craig | Iliamna | Livengood | Pitkas Point | Teller |
| Alakanuk | Crooked Creek | Ivanof Bay | Lower Kalskag | Platinum | Tenakee Springs |
| Alcan Border | Eagle | Kake | Lutak | Point Baker | Tetlin |
| Aleknagik | Eagle Village | Kaktovik | Manokotak | Pope-Vannoy Landing | Thorne Bay |
| Aleneva | Edna Bay | Kaltag | Mekoryuk | Port Alexander | Toksook Bay |
| Allakaket | Elfin Cove | Karluk | Metlakatla | Port Alsworth | Tununak |
| Anaktuvuk Pass | Elim | Kasaan | Mosquito Lake | Port Heiden | Twin Hills |
| Aniak | Eureka Roadhouse | Kasigluk | Mud Bay | Port Lions | Unalakleet |
| Arctic Village | Excursion Inlet | King Cove | Naknek | Port Protection | Unalaska |
| Atka | False Pass | King Salmon | Naukati Bay | Portage Creek | Upper Kalskag |
| Atmautluak | Fort Yukon | Kipnuk | Nelson Lagoon | Red Devil | Venetie |
| Atqasuk | Gambell | Kivalina | Newhalen | Sand Point | Wales |
| Beluga | Game Creek | Klawock | Newtok | Savoonga | White Mountain |
| Birch Creek | Glacier View | Kodiak City | Nightmute | Shaktoolik | Whitestone Logging |
| Brevig Mission | Goodnews Bay | Kodiak Station | Nikolski | Sitka | Camp |
| Central | Gustavus | Kokhanok | Nome | Skagway | Whittier |
| Chalkyitsik | Healy Lake | Kongiganak | Nondalton | Sleetmute | Womens Bay |
| Chefornak | Hobart Bay | Kotlik | Nunam Iqua | South Naknek | Wrangell |
| Chevak | Hollis | Koyuk | Nunapitchuk | St. George | |
| Chicken | Holy Cross | Koyukuk | Old Harbor | St. Marys | |
| Chiniak | Hooper Bay | Kupreanof city | Ouzinkie | St. Michael | |
| Chuathbaluk | Huslia | Kwigillingok | Pedro Bay | St. Paul | |
| Circle | | Lake Minchumina | | Stony River | |
| Coffman Cove | | | | | |
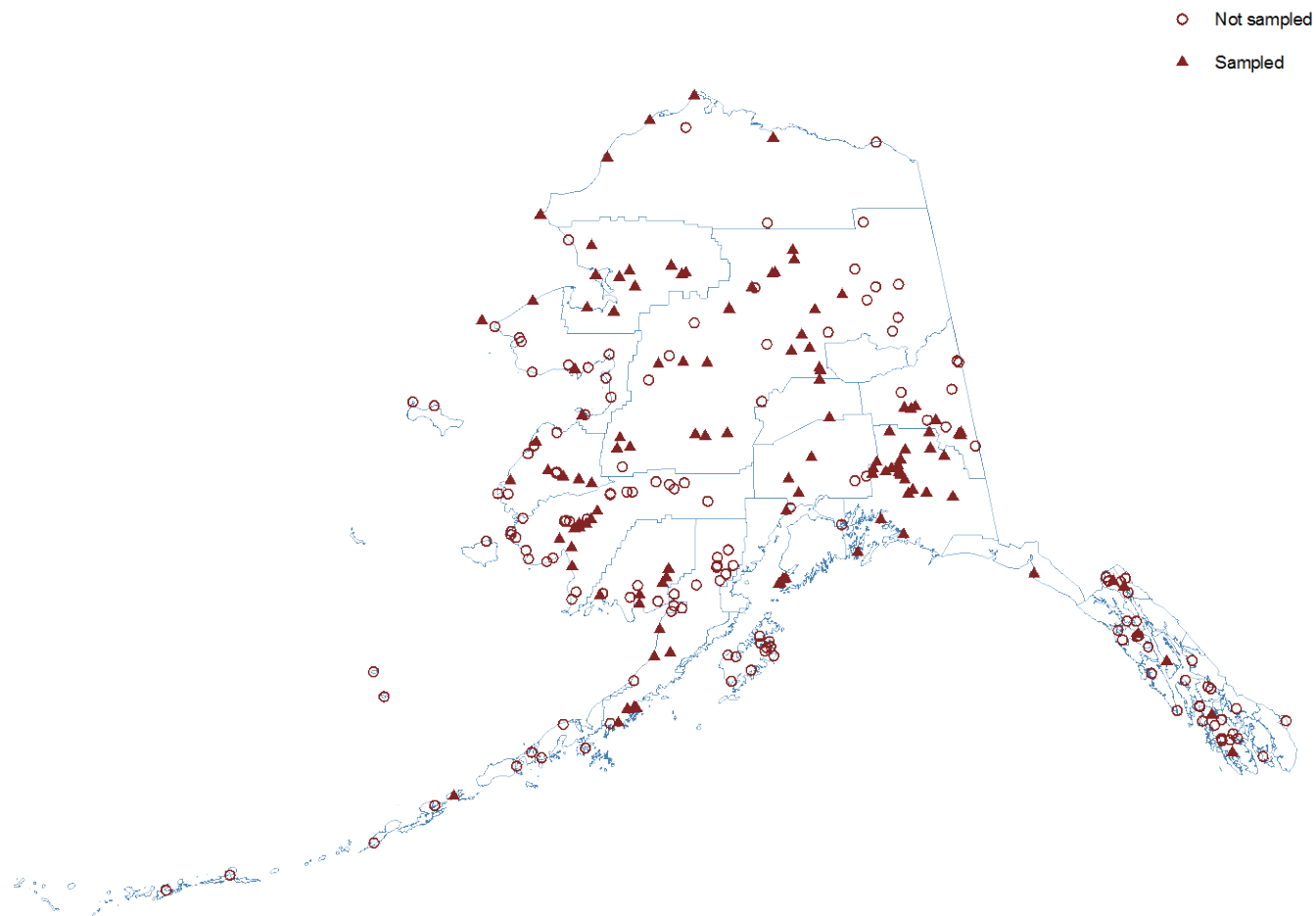
**Figure C1. Location of Sampled and Non-Sampled Communities.** [56]

---

[56] The following description of the boundaries in the figure are from the census site https://www.census.gov/geo/reference/gtc/gtc_cou.html, accessed 12/17/18. "In Alaska, …[geographic] entities are the organized boroughs, city and boroughs, municipalities, and census areas; the latter of which are delineated cooperatively for statistical purposes by the state of Alaska and the Census Bureau." The map was created using the R software module Tigris.

**Figure C2. Alaska and its six regions.**

# Appendix D. Calculation of the Mean and the 90th Percentiles of the Resource Use Rates

References cited in this appendix can be found in the References section of the main report (section 6).

TMWL calculated consumption means and percentiles using standard survey analysis methods. Specifically, the aggregated dataset was set up as a post-stratified one-level cluster sampling design, where the cluster was the combination of the community and stratum (a total of 113 clusters).[57] The one-level cluster sampling design includes two sets of weights: cluster (community) weights and observation (household) weights. In an earlier, exploratory analysis (not shown) TMWL set the cluster weights to 1.0 (an equal weight for every community).[58] However, note that the varying size (number of households) of the communities also causes the number of observations to vary (and their weights) in each community (described next). Secondly, the varying proportion of communities sampled per region are adjusted later by post-stratifying (Levy and Lemeshow 1999) on the region (see below). The observation weights are based on the survey's target sample size for a community, the non-response rate within the community, and the size of each household. The size of the household is the number of individuals that each household observation represents in that community cluster. The observation weights were calculated as follows:

- Non-response adjusted sampling household weights = # households in the community / # household from the community in the analysis set. The analysis set includes all households which responded to the survey and had useable data.
- Observation weights = non-response adjusted sampling household weights * household size.

The final weights are the product of the cluster (community) weights and the observation (household) weights with an additional post-stratification adjustment for the region. Without

---

[57] The number of clusters (113) is different from the number of communities in the harvest data (118) for two reasons. **First**, strata within communities were treated as clusters. Community #113, Dillingham, had two strata and community #185, Kenny Lake, had three strata. Each stratum for these two communities was treated as a separate "community" for analysis. In this collection of community surveys, the strata variable in the ADF&G dataset had a different meaning compared to how strata are defined in standard survey weighting (i.e., the ADF&G strata variable had a different meaning across different communities, while "strata" in survey analysis has a common meaning across the entire population). As a reasonable (and convenient) solution we simply treated the strata as separate clusters. As this choice impacts the handling of only two communities, we do not expect it to impact the estimated means and percentiles in a major way. **Second**, seven communities (clusters) in the harvest data included multiple (two or three) communities: Dot Lake Village (community #495) was included in Dot Lake (community #115), Four Mile Road (community #1003) was included in Nenana (community #241), Northway Junction (community #437) and Northway Village (community #438) were included in Northway (community #256), Willow Creek (community #491) was included in Kenny Lake (community #185), Seldovia Village (community #573) was included in Seldovia (community #304), Nabesna (community #235) was included in Slana (community #316), Silver Springs (community #785) was included in Copper Center (community #103). In summary, eight communities were merged (as just noted) with other communities and two communities were subdivided (by strata) to add three "communities" (clusters) for the analysis so that $118 - 8 + 3 = 113$ clusters.

[58] The key for the noted earlier analysis is that the community weights are all equal. It does not matter that they are set to a specific positive constant (in this case 1.0 for the earlier, exploratory analysis).

post-stratification, the sum of the weights[59] would quantify the total number of individuals that all the surveyed households represent in the 118 sampled communities. The post-stratification adjusts these weights by a scaling factor to ensure that they represent all relevant communities (N = 262) and to account for factors that may influence the tendency of a community to be sampled by ADF&G. Because the region of the community was the most prominent factor in the tendency of a community to be sampled (see Table D1), post-stratification was based on region alone.[60] In the post-stratification calculation, one must specify the true totals of individuals within each post-stratum (see the second numeric column in Table D1). These totals imply specific scaling factors for the weights (see the third numeric column). For example, before the post-stratification, the sample of communities in the Southeast region represented 4,038 individuals. As the true total population is 25,328 (according to the 2010 U.S. Census) all weights had to be multiplied by a factor of 6.272.[61]

**Table D1. Regional Adjustment Factor for the Weights**

| Region | The estimated total population in the communities sampled by ADF&G* | Total population** | Regional adjustment factor for the weights |
|---|---|---|---|
| Southeast | 4,038 | 25,328 | 6.272 |
| Southcentral | 6,692 | 7,183 | 1.073 |
| Southwest | 5,103 | 20,438 | 4.005 |
| Western | 13,151 | 24,038 | 1.828 |
| Arctic | 15,353 | 22,979 | 1.497 |
| Interior | 4,798 | 7,501 | 1.563 |

Notes: This table is a duplicate of Table 1 of the report.
*Equals the sum of weights before post-stratification (i.e., the participation factor x the household size)
**Total population living in households in rural communities and excluding population residing in group quarters (U.S. Census 2010)

Within the weighting scheme described above, TMWL estimated means and percentiles of the use levels using standard survey analysis methods as implemented in the R software package "survey" (R Core Team 2017, Lumley 2004, Lumley 2019). Variance statistics (standard errors and 95% confidence intervals) for the estimated means and percentiles were also obtained using standard methods: linearization for means, and the score method for percentiles (Binder 1991). To assess the statistical significance of the differences in mean use levels across different regions, TMWL compared two survey linear regression models: (1) the use levels regressed on the intercept only[62] and (2) the survey linear regression model with the use levels regressed on the intercept <u>and</u> variables representing the six regions of Alaska.[63] This comparison addresses

---

[59] The product of the cluster (community) weights and the observation (household) weights.
[60] Specifically, it is based on the total number of individuals living in households in each region, based on the U.S. 2010 Census.
[61] Note that TMWL's calculations specify that the post-stratification is done with the specific population totals. This is important to ensure that correct confidence intervals can be calculated.
[62] This model corresponds to the null hypothesis that the means for all regions are equal.
[63] This model corresponds to the alternative hypotheses that the means for all regions are not equal.

the question of whether adding the regions to the regression model improves the fit of the model, taking potential random improvement into account. The statistical significance of the improvement is assessed using the Rao-Scott likelihood ratio test.[64] To assess the statistical significance of the differences among the six regions in percentiles of use rates (i.e., 90th percentile in this report), TMWL used the standard Wald test (with the test statistics based on the point estimates and the region-specific variances described above).

Table D2, offered for completeness, shows the number of households providing survey data for the rate calculations in this report. The number of household members is also shown, though data were available only per household and not per household member.

**Table D2. Number of Responding Households and Number of Household Members Covered by ADF&G Interview Data Used in This Report. Includes Only Households Providing Useable Data for the Calculation of Resource Use Rates.**

| Region | Number of households | Number of household members |
|---|---:|---:|
| Southeast | 499 | 1,246 |
| Southcentral | 1,218 | 3,196 |
| Southwest | 645 | 2,122 |
| Western | 1,550 | 6,504 |
| Arctic | 1,663 | 6,619 |
| Interior | 1,057 | 2,751 |
| All regions | 6,632 | 22,438 |

---

[64] The Rao-Scott likelihood ratio test can be thought of as the weighted equivalent to the one-way ANOVA test for unweighted data.