Special Publication No. SP 09-08

Sampling Considerations for Estimating Geographic Origins of Chinook Salmon Bycatch in the Bering Sea Pollock Fishery

by

Jerome J. Pella

and

Harold J. Geiger

May 2009

Alaska Department of Fish and Game Divisions of Sport Fish and Commercial Fisheries

Symbols and Abbreviations

The following symbols and abbreviations, and others approved for the Système International d'Unités (SI), are used without definition in the following reports by the Divisions of Sport Fish and of Commercial Fisheries: Fishery Manuscripts, Fishery Data Series Reports, Fishery Management Reports, and Special Publications. All others, including deviations from definitions listed below, are noted in the text at first mention, as well as in the titles or footnotes of tables, and in figure or figure captions.

Weights and measures (metric)

centimeter	cm
deciliter	dL
gram	g
hectare	ha
kilogram	kg
kilometer	km
liter	L
meter	m
milliliter	mL
millimeter	mm

Weights and measures (English)

cubic feet per second	ft^3/s
foot	ft
gallon	gal
inch	in
mile	mi
nautical mile	nmi
ounce	oz
pound	lb
quart	qt
yard	yd

Time and temperature

day	d
degrees Celsius	°C
degrees Fahrenheit	°F
degrees kelvin	K
hour	h
minute	min
second	s

Physics and chemistry

all atomic symbols	
alternating current	AC
ampere	A
calorie	cal
direct current	DC
hertz	Hz
horsepower	hp
hydrogen ion activity (negative log of)	pH
parts per million	ppm
parts per thousand	ppt, ‰
volts	V
watts	W

General

Alaska Administrative Code	AAC
all commonly accepted abbreviations	e.g., Mr., Mrs., AM, PM, etc.
all commonly accepted professional titles	e.g., Dr., Ph.D., R.N., etc.
at	@
compass directions:	
east	E
north	N
south	S
west	W
copyright	©
corporate suffixes:	
Company	Co.
Corporation	Corp.
Incorporated	Inc.
Limited	Ltd.
District of Columbia	D.C.
et alii (and others)	et al.
et cetera (and so forth)	etc.
exempli gratia (for example)	e.g.
Federal Information Code	FIC
id est (that is)	i.e.
latitude or longitude	lat. or long.
monetary symbols (U.S.)	$, ¢
months (tables and figures): first three letters	Jan,...,Dec
registered trademark	®
trademark	™
United States (adjective)	U.S.
United States of America (noun)	USA
U.S.C.	United States Code
U.S. state	use two-letter abbreviations (e.g., AK, WA)

Measures (fisheries)

fork length	FL
mideye to fork	MEF
mideye to tail fork	METF
standard length	SL
total length	TL

Mathematics, statistics

all standard mathematical signs, symbols and abbreviations	
alternate hypothesis	H_A
base of natural logarithm	*e*
catch per unit effort	CPUE
coefficient of variation	CV
common test statistics	(F, t, χ^2, etc.)
confidence interval	CI
correlation coefficient (multiple)	R
correlation coefficient (simple)	r
covariance	cov
degree (angular)	°
degrees of freedom	df
expected value	*E*
greater than	>
greater than or equal to	≥
harvest per unit effort	HPUE
less than	<
less than or equal to	≤
logarithm (natural)	ln
logarithm (base 10)	log
logarithm (specify base)	log_2, etc.
minute (angular)	'
not significant	NS
null hypothesis	H_O
percent	%
probability	P
probability of a type I error (rejection of the null hypothesis when true)	α
probability of a type II error (acceptance of the null hypothesis when false)	β
second (angular)	"
standard deviation	SD
standard error	SE
variance	
population	Var
sample	var

SPECIAL PUBLICATION NO. 09-08

SAMPLING CONSIDERATIONS FOR ESTIMATING GEOGRAPHIC ORIGINS OF CHINOOK SALMON BYCATCH IN THE BERING SEA POLLOCK FISHERY

By
Jerome J. Pella
and
Harold J. Geiger
St. Hubert Research Group, Juneau, AK 99801

SHRG@ak.net

907.723.3234

Alaska Department of Fish and Game
Division of Sport Fish, Research and Technical Services
333 Raspberry Road, Anchorage, Alaska, 99518-1565

May 2009

This investigation was primarily financed by the Bering Sea pollock cooperatives and in part by the Alaska Department of Fish and Game. This paper was completed under contract with the Alaska Department of Fish and Game. Views in this paper are those of the authors and do not represent the State of Alaska.

The Special Publication series was established by the Division of Sport Fish in 1991 for the publication of techniques and procedures manuals, informational pamphlets, special subject reports to decision-making bodies, symposia and workshop proceedings, application software documentation, in-house lectures, and became a joint divisional series in 2004 with the Division of Commercial Fisheries. Special Publications are intended for fishery and other technical professionals. Special Publications are available through the Alaska State Library, Alaska Resources Library and Information Services (ARLIS) and on the Internet: http://www.sf.adfg.state.ak.us/statewide/divreports/html/intersearch.cfm. This publication has undergone editorial and peer review.

Jerome J. Pella,
St. Hubert Research Group, 222 Seward Street, Suite 205, Juneau, AK 99801
USA

Harold J. Geiger,
St. Hubert Research Group, 222 Seward Street, Suite 205, Juneau, AK 99801
USA

This document should be cited as:
Pella, J. J., and H. J. Geiger. 2009. Sampling considerations for estimating geographic origins of Chinook salmon bycatch in the Bering Sea pollock fishery. Alaska Department of Fish and Game, Special Publication No. 09-08, Anchorage.

The Alaska Department of Fish and Game (ADF&G) administers all programs and activities free from discrimination based on race, color, national origin, age, sex, religion, marital status, pregnancy, parenthood, or disability. The department administers all programs and activities in compliance with Title VI of the Civil Rights Act of 1964, Section 504 of the Rehabilitation Act of 1973, Title II of the Americans with Disabilities Act (ADA) of 1990, the Age Discrimination Act of 1975, and Title IX of the Education Amendments of 1972.

If you believe you have been discriminated against in any program, activity, or facility please write:

ADF&G ADA Coordinator, P.O. Box 115526, Juneau, AK 99811-5526

U.S. Fish and Wildlife Service, 4401 N. Fairfax Drive, MS 2042, Arlington, VA 22203

Office of Equal Opportunity, U.S. Department of the Interior, 1849 C Street NW MS 5230, Washington DC 20240

The department's ADA Coordinator can be reached via phone at the following numbers:

(VOICE) 907-465-6077, (Statewide Telecommunication Device for the Deaf) 1-800-478-3648, (Juneau TDD) 907-465-3646, or (FAX) 907-465-6078

For information on alternative formats and questions on this publication, please contact:

ADF&G Division of Sport Fish, Research and Technical Services, 333 Raspberry Road, Anchorage AK 99518 (907) 267-2375.

TABLE OF CONTENTS

Page

LIST OF TABLES ii
LIST OF FIGURES ii
LIST OF APPENDICES ii
PREFACE iii
ABSTRACT 1
EXECUTIVE SUMMARY 1
INTRODUCTION AND BACKGROUND 4
The Fisheries 5
The Bycatch 7
The Observer Program 9
Processors, fishing vessels, and their coverage and information through the Observer Program 10
Sampling methods on trawlers 12
The Fisheries Monitoring and Analysis Division databases as of January 2008 12
SPECIFIC SAMPLING TERMS AND ESTIMATION ISSUES 13
The Current Approach to Estimating a Bycatch Total 15
Alternate Information About Salmon Bycatch 16
Important Tests of Assumptions 17
Improved Bycatch Estimates With Incomplete Information 18
Sample Selection of Individual Specimens for Stock and Age Composition Estimates 20
GENETICS-BASED INFORMATION ON STOCK ORIGINS OF THE BYCATCH 23
Uncertainty and Sample Size 27
A Simple Example Mixture Analysis Using Colored Balls and Urns 28
RECOMMENDATIONS 30
Estimation of Domain Bycatch 30
Genetic Analysis 32
Fishery samples 32
Baseline evaluation 33
Bycatch stock mixture analysis 35
Key Recommendations 35
CONCLUSION 36
ACKNOWLEDGEMENTS 37
REFERENCES CITED 38
APPENDICES 43

LIST OF TABLES

Table **Page**

1. Catch and bycatch statistics from the 2007 Bering Sea pollock fishery, by vessel category.....7
2. Estimated salmon bycatch in Bering Sea–Aleutian Islands directed pollock fisheries in thousands of fish. The estimates include catches in the Community Development Quota sector of the fishery.....9
3. Contractors with permits from the National Marine Fisheries Service to provide observers for the North Pacific Groundfish Observer Program. 10
4. Numbers of colored balls in 3 source urns and the combined mixture urn. 28

LIST OF FIGURES

Figure **Page**

1. Major fishing regions and important river systems of Alaska.....6

LIST OF APPENDICES

Appendix **Page**

A. Observer Forms and Information Recorded on Trawlers and at Plants.....44
B. Estimators, Statistical Summaries, and Reported Information When Observers Cannot Census the Bycatch—Miller's Methods 45
C. The General Mixture Problem and its Solution.....52

PREFACE

In February of 2008, salmon bycatch in the Bering Sea/Aleutian Islands pollock fishery was once again identified as a primary concern by the North Pacific Fishery Management Council (Council). At its February meeting, the Council forwarded a problem statement for analysis that included the following language:

> "*the State of Alaska has been forced to close or greatly reduce some commercial, subsistence and sport fisheries in Western Alaska. Reasons for reductions in the number of Chinook salmon returning to spawn in Western Alaska rivers and the Canadian portion of the Yukon River drainage are uncertain, but recent increases in Bering Sea bycatch may be a contributing factor*".

The Council recognized that current salmon sampling protocols were not adequate to fully understand the effects of Bering Sea bycatch on Pacific Chinook salmon stocks. A more thorough understanding of geographic origin and age distribution of salmon bycatch is needed to complement management actions focused on reducing bycatch.

The Alaska Department of Fish and Game (ADF&G), using funds largely provided by the Bering Sea pollock industry, requested proposals to design a robust sampling plan to estimate bycatch stock of origin using genetic stock identification methods. The specific goal was to develop a sampling plan and corresponding statistical estimators of stock composition and stock abundance for Chinook and chum salmon caught by vessels in the Bering Sea pollock fishery.

The paper *Sampling Considerations for Estimating Geographic Origins of Chinook Salmon Bycatch in the Bering Sea Pollock Fishery* completes a contract established with the St. Hubert Research Group of Juneau. The report outlines fundamental aspects of a defensible sampling program and provides recommendations that may serve as a framework for improved sampling protocols in the future. The authors have expertise in scientific sampling and genetic mixture analysis. Information provided in the report regarding the North Pacific Groundfish Observer Program and Catch Accounting System has not been fully reviewed by the National Marine Fisheries Service.

The authors are not experts on these federal programs. We recognize that designing and implementing a robust sampling plan that meshes with the groundfish observer program will ultimately rely on the National Marine Fisheries Service's expertise. We look forward to their review and strongly support cooperative efforts between the Council, the National Marine Fisheries Service, and ADF&G to achieve the shared goal of obtaining demographic information of salmon bycatch.

ABSTRACT

Chinook salmon bycatch in the Bering Sea–Aleutian Islands pollock fishery steadily increased from a low of 5,000 fish in 2000 to over 121,000 fish in 2007. This led fishery managers to call for improved stock-specific estimates of the bycatch. Statistical estimation of stocks of origin of the bycatch will require (1) reliable estimates of the bycatch magnitude for temporal and spatial divisions of the pollock fishery; (2) representative collections of Chinook salmon tissues from these temporal and spatial divisions, suitable for biochemical genetic analysis; and (3) suitable Chinook salmon baseline samples and genetic mixture algorithms. Fortunately, the National Marine Fishery Service has a well-developed observer program that can be used as a platform to develop a program for genetic stock mixture analysis of Chinook salmon in the bycatch, although this program will need additional resources. Before moving forward, we recommend that the National Marine Fisheries Service improve their methods for estimating the bycatch magnitude, or at least that they provide improved descriptions of their methods. We also recommend developing strict sampling protocols that ensure representative sampling of Chinook salmon in the bycatch, and we are especially concerned that sampling not be size selective and that sampling rates be as uniform as possible across hauls, deliveries, vessels, and so on. To achieve these objectives, we recommend that the National Marine Fisheries Service Groundfish Observer Program systematically sample the Chinook salmon bycatch for tissue samples that are used for genetic analysis. Finally, we have detailed recommendations for a Bayesian statistical approach to the actual mixture analysis.

Keywords: Chinook salmon, bycatch, Bering Sea–Aleutian Islands, pollock, observer program, stock mixture analysis.

EXECUTIVE SUMMARY

From 2000 to 2007, the bycatch of Chinook salmon in the Bering Sea–Aleutian Islands pollock fisheries strongly trended upwards (although the bycatch dropped in 2008). In 2007, the year with the highest bycatch, this pollock fishery took 1.32 million metric tons of pollock with a bycatch of over 121,000 Chinook salmon. Responding to the bycatch increase, fishery managers concluded that they now need an assessment of its effect on the contributing Chinook salmon stocks that are already fully allocated. Also, fishery managers concluded that they need a way to predict how measures to control the bycatch may affect other commercial, subsistence, and sport fisheries for Chinook salmon in Alaska and in Canada. Here we describe a sampling plan and associated statistical analysis that will produce stock-specific harvest estimates for Chinook salmon in this bycatch. In turn, fishery managers can use these harvest estimates to assess the annual effects of the pollock fishery on individual inshore runs of Chinook salmon.

In the 1970s, the National Marine Fisheries Service and fishing industry together started an observer program to monitor the pollock fishery. The current form of the observer program was established by 1990, and this program now has several hundred technicians stationed on vessels and at onshore processing stations throughout the pollock fishery. Because this is a mature sampling program that performs many essential tasks, the importance of enumerating and sampling Chinook salmon in the bycatch must be weighed against the other competing assignments unless added resources are provided. Even so, this program provides the obvious platform from which to build a system of data and specimen collection by which to develop stock-of-origin estimates of the Chinook salmon bycatch.

Complicating the sampling of Chinook salmon bycatch in the pollock fishery, these salmon occur throughout catches of the distinct inshore, offshore, mothership, and Community Development Quota sectors of the fishery. Further, the bycatch is caught by many individual trawl vessels: larger catcher-processor vessels that both fish and process the catch, as well as by smaller catcher vessels that fish and deliver catch to processors in any of the sectors. Catcher-processor vessels, as well as motherships that only process fish delivered from catcher vessels,

have observers onboard to monitor most, but not necessarily all of the catch being processed. Inshore processing plants that receive fish from catcher vessels have observer coverage that depends on the amount of groundfish processed by them each month: plants processing 1,000 metric tons or more must have an observer present each day of processing, and plants processing between 500 metric tons and 1,000 metric tons must have an observer present for 30% of those days. Catcher vessels deliver primarily to the inshore sector and are divided into three classes based on their lengths, and observer coverage requirements depend on the class. Small catcher vessels less than 60′ that deliver groundfish are exempt from observer coverage, sampling station, and scale requirements. However, vessels in this size class generally fish in the Gulf of Alaska and none currently participate in the Bering Sea pollock fishery. Midsized catcher vessels from 60′ to less than 125′ are required to have an observer aboard for at least 30% of fishing days each quarter. The captain of each such vessel selects the days to be sampled. Large catcher vessels 125′ and greater are required to have an observer aboard 100% of fishing days. The upshot of the observer program coverage requirements is that the vast majority of hauls by the trawl fleet are monitored directly at sea or at offloading locations aboard motherships or shore-based processing plants. The practical challenge associated with getting a representative sample of tissues for genetic analysis from the Chinook salmon bycatch has to do with physically getting access to all of the fish irrespective of sector, haul, vessel, or processing destination.

If a random sample of Chinook salmon tissues that is representative of the stock composition can be collected from some subdivision of the Chinook salmon bycatch, biochemical genetic techniques and statistical tools can be used to produce the proportional estimates of stock origins. New biochemical techniques are constantly under development, and the existing techniques are continuously being improved. Briefly, these techniques all depend on the development of a profile of genetic types from a sample of specimens from each of the populations potentially contributing to the bycatch mixture. These samples from which the genetic profiles of the contributing stocks are derived are collectively called *the baseline*. Two baselines have been developed for Chinook salmon using microsatellites and single nucleotide polymorphisms (SNPs) of nuclear DNA, and these could be used separately when their geographical coverage is considered adequate, or even in combination for greater stock resolving power. Next, the genetic types from a representative bycatch sample are compared statistically to the baseline profiles to decide on their probable origins using *genetic mixture analysis*. Although there are several algorithms in use to perform genetic mixture analysis, we recommend a Bayesian method that extracts the information from the genetic types seen in the fishery sample to refine the genetic profiles from the baseline. Further, we recommend that the genetic mixture analysis be validated using simulations run on the genetic baseline. Commonly, so-called 100% mixtures are used in these simulation studies. The 100% mixtures in the simulations likely provide an optimistic view of performance of mixture analysis. In addition, recent studies of the standard simulation method shows that it may overstate the predicted accuracy with profound bias for closely related stocks. We recommend cross validation methods and simulation of "worst case" equal-contribution mixtures for these validation studies. Once these simulation studies demonstrate the practicality of the method and the baseline, the final products of the genetic mixture analysis are estimates of the proportional contributions to the bycatch mixture by regional baseline stock groups, and possibly even finer stock resolution.

These proportional estimates of stock origins can be converted into estimates of magnitude (i.e., number of fish) by multiplication with the number of Chinook salmon in the corresponding

bycatch. Although the National Marine Fisheries Service has estimated the number of Chinook salmon in the bycatch for many years, their methods have not been carefully documented. Bycatch numbers could, in principle, be ascertained by what is commonly called a census: an essentially complete enumeration, or count, of each Chinook salmon in the bycatch. This approach is simple, easy to explain, and has the advantage that it is free of sampling error, i.e., error that results from examining and counting only a portion of the larger statistical population rather than its entirety. Currently, regulations are in place requiring that each Chinook salmon in the bycatch be turned into a fishery observer before it is discarded, although the National Marine Fisheries Service has not used these regulations to develop a census program. If the National Marine Fisheries Service adopts this census approach, simple statistical tests can be used to detect and discourage misrepresentation or misreporting of the bycatch in the fleet as a whole. However, if this census approach is not acceptable, sound statistical methods by which to estimate Chinook salmon bycatch in this pollock fishery, are carefully documented in a PhD thesis from the University of Washington. Currently, the National Marine Fisheries Service is not using these methods either. We have abstracted key equations from this thesis, and provided them in an appendix. If neither of these options is acceptable, we strongly recommend that the National Marine Fisheries Service provide detailed descriptions of their methods, and we recommend that these methods be peer reviewed and published in the scientific literature.

Returning to the issue of sample collections of tissues for genetic analysis, a failure to have a sample of sufficient size from the bycatch that is also representative of its stock composition will have an unpredictable effect on the apparent origins of the bycatch. For that reason, we recommend that sampling of Chinook salmon in the bycatch be fully funded, and we strongly recommend that all specimens that are analyzed for mixture analysis be collected under strict protocols that ensure the Chinook salmon stocks present are represented in proportion to their occurrence. Because fish age is almost certainly associated with stock of origin for probably any temporal and spatial division in the Bering Sea, the sampling process must be strictly non-selective for size or age. Because fish could mechanically sort by size within a trawl, protocols should be put into place to ensure that the sampling process selects specimens throughout each haul or delivery at nearly the same rate. Because stocks may tend to cluster within the harvest, the sampling must spread tissue collections broadly among vessels, hauls, and deliveries with the sampling rate held constant. While we did not make a recommendation for a specific sample size, we do indicate fairly routine statistical methods by which to determine it. Although larger sample sizes will result in more precise estimates, the available financial resources will probably determine the final sample sizes each year. Even so, under a wide range of assumptions, the genetic mixture analysis will require, at a minimum, several hundreds of tissue specimens for each time and area domain for which geneticists estimate stock composition of the Chinook salmon bycatch.

To guarantee that the tissue samples are representative of the stock composition of the bycatch, we recommend that tissue specimens be obtained through a *systematic sampling* of the bycatch, with the sampling rate fixed across sampling staff. By that we mean that the sampling technicians should go through the bycatch, one Chinook salmon at a time, and that a tissue specimen be collected from every n^{th}, fish, with n to be determined before the season. By fixing this systematic sampling rate (the systematic sampling rate will be $1/n$), the individual tissue specimens can be flexibly grouped into arbitrary time and area domains of interest, so as to allow the most efficient use of laboratory resources. Additionally, this systematic sampling will be

operationally simple, efficient, and it will avoid expensive culling of specimens in order to standardize sampling rates that might have occurred under alternative sampling schemes.

We developed this plan assuming that the goal was to estimate the Chinook salmon bycatch by its regional geographic origins for the entire Bering Sea–Aleutian Islands trawl fishery. We also assumed that the responsible agencies would receive the resources that they need for representative sampling of the entire Chinook salmon bycatch. However, if adequate resources cannot be secured to collect and process with the best available techniques a fully representative sample of sufficient size from the entire Chinook bycatch, then we recommend that the goals of this larger effort be scaled back so that more limited targets match the resources that are actually available.

INTRODUCTION AND BACKGROUND

The term *bycatch* is typically used to describe those fish caught, whether utilized or not, while fishing for some other species. Pacific salmon, steelhead trout, herring, halibut, and other species are considered bycatch and *prohibited species* in the Bering Sea–Aleutian Islands and Gulf of Alaska groundfish fisheries. Except for vessels permitted to participate in a program for donation to food banks, these species cannot be retained and used by a vessel owner, irrespective of whether the prohibited species of fish can be returned to the water alive. Because a large fraction of the Chinook salmon in the Bering Sea–Aleutian Islands pollock fishery bycatch apparently originate from Western Alaska stocks, including the Yukon River (Myers and Rogers 1988), and because these stocks of Chinook salmon are fully allocated and utilized (e.g., United States and Canada Yukon River Joint Technical Committee 2007), their incidental harvest is a matter of serious concern to the inshore salmon fishers. Also, because many of the people that live along the Yukon River have few economic opportunities, the salmon resource takes on a heightened significance in the region. Understanding the effects that the groundfish fisheries are having on the inshore abundance of fully utilized Pacific salmon stocks has taken on a sense of urgency propelled in part by recent major news stories in Alaska and elsewhere about the groundfish bycatch. The North Pacific Fishery Management Council has been especially concerned by the Chinook salmon bycatch in the Bering Sea for years, and has continued to evolve regulatory measures intended to control it to the extent possible while enabling an efficient pollock fishery (Anonymous 2008).

A necessary first step in understanding and responding to this problem is to assess the magnitude and geographic sources of the bycatch. Here we outline and describe a plan for estimating the annual Chinook salmon bycatch in the entire Bering Sea–Aleutian Islands pollock fishery. Naturally, this program will require additional resources to put into operation. Fortunately, a well-developed monitoring program of the fishery has been in place for two decades and continues to improve. The data from this program are invaluable for determining the magnitude of bycatch. Fortunately again, we can now make use of biochemical genotyping and statistical tools that permit ever-finer resolution of catch mixtures of Pacific salmon to stocks of origin (e.g., Shaklee et al. 1999; Pella and Masuda 2001; and many others). However, no matter how refined the laboratory methods, or how complex the statistical algorithms, these tools will not be helpful unless they are used as a part of a well-designed study of the issue. Our intent is to first examine the fisheries and their history, review the written descriptions of the current National Marine Fisheries Service fishery monitoring program and interview staff familiar with its operations, cover some pertinent terms and concepts, review the mathematics and sampling

models underlying genetic population mixture analysis, and conclude with recommendations for data collection and analysis related to estimation of the magnitude of the Chinook salmon bycatch and its regional geographic stock composition.

We will only directly address the bycatch issue for Chinook salmon, but the principles involved in estimating the chum salmon bycatch are exactly the same. So without loss of generality, we will leave it to the reader concerned with chum salmon bycatch to mentally replace "Chinook" with "chum" where appropriate throughout the document. An exception occurs in the section "Genetics-Based Information on Stock Origins of the Bycatch" where Chinook salmon and chum salmon genetic baselines are described separately.

THE FISHERIES

To the uninitiated, these groundfish fisheries in the Bering Sea–Aleutian Islands and the Gulf of Alaska constitute a confusing mix of directed fisheries for walleye pollock, yellowfin sole, Pacific cod, Atka mackerel, Rock sole, Pacific halibut, and many other species. In trying to understand the fishery statistics of the region, sometimes the reporting is for the Bering Sea–Aleutian Islands fisheries, sometimes these two regions are reported separately, and sometimes the statistics include the Gulf of Alaska fisheries (Figure 1). In describing prohibited-species bycatch, at times the statistics include just the pollock fisheries, and still other times they include all groundfish fisheries together. We mention this as a word of caution when comparing statistics from different sources. What seem to be substantial differences in magnitude often represent differences in reporting category or fishery description.

The walleye pollock species (*Theragra chalcogramma*) currently makes up the most valuable resource in these groundfish fisheries (Woodby et al. 2005). Over 95% of the pollock harvest in the Bering Sea and Aleutian Islands fisheries are taken with pelagic trawls. Indeed, pelagic trawls are the only legal gear for directed pollock fishing in this area. In the Gulf of Alaska fishery, 90% of the harvest is taken with trawls (Woodby et al. 2005). The problem of salmon bycatch is almost exclusively a problem with the trawl gear type. In the 1990s, over 99% of the salmon bycatch was taken in trawl gears. The reported bycatch from the Gulf of Alaska has been lower (both for Chinook salmon and for "other" salmon) than it has been for the Bering Sea–Aleutian Islands area, and the bycatch in the Gulf has drawn less attention (Witherell et al. 2002). The term "other" salmon refers to the aggregate of chum, coho, sockeye, and pink salmon, but most (roughly 95%) individuals in the bycatch are chum salmon.

Foreign fleets dominated commercial fishing in the Bering Sea from the early 1930s until after passage of the Magnuson Fishery Conservation and Management Act in 1976, and, over time, these fishing fleets were converted into U.S.-foreign joint ventures, which then finally were replaced by the current fleet of U.S. catchers (Megrey and Wespestad 1990; Witherell and Pautzke 1997). Under the Magnuson Act, eight regional councils were established to manage U.S. fisheries, including the North Pacific Fishery Management Council, whose jurisdiction is the 900,000 square mile Exclusive Economic Zone off the Alaska coast. To encourage investment in the fishery, the American Fisheries Promotion Act was passed in 1980. Under the Promotion Act, fish quotas were given preferentially to nations that aided the development of the U.S. fishing industry. By 1991, all foreign commercial fishing within the Exclusive Economic Zone had been displaced by domestic fishing. In 1992, vessels under the Western Alaska Community Development Quota Program first began fishing pollock from the Bering Sea–Aleutian Islands area of the Alaska Exclusive Economic Zone. The goal of the program was to

promote fisheries-related economic development in Western Alaska. A Community Development Quota reserve of 7.5% of the *Total Allowable Catch* for pollock in the Bering Sea–Aleutian Islands fishery was allocated to community organizations with an approved plan. The fish could be harvested by the organizations themselves, or they could contract a non-Community Development Quota firm to harvest the fish. The fishing could occur after the open access quota had been caught and that fishery closed. The Magnuson Act was re-authorized in 1996 as the Magnuson-Stevens Fishery Conservation and Management Act, by which the U.S. declared management authority over all fish resources within its Exclusive Economic Zone.

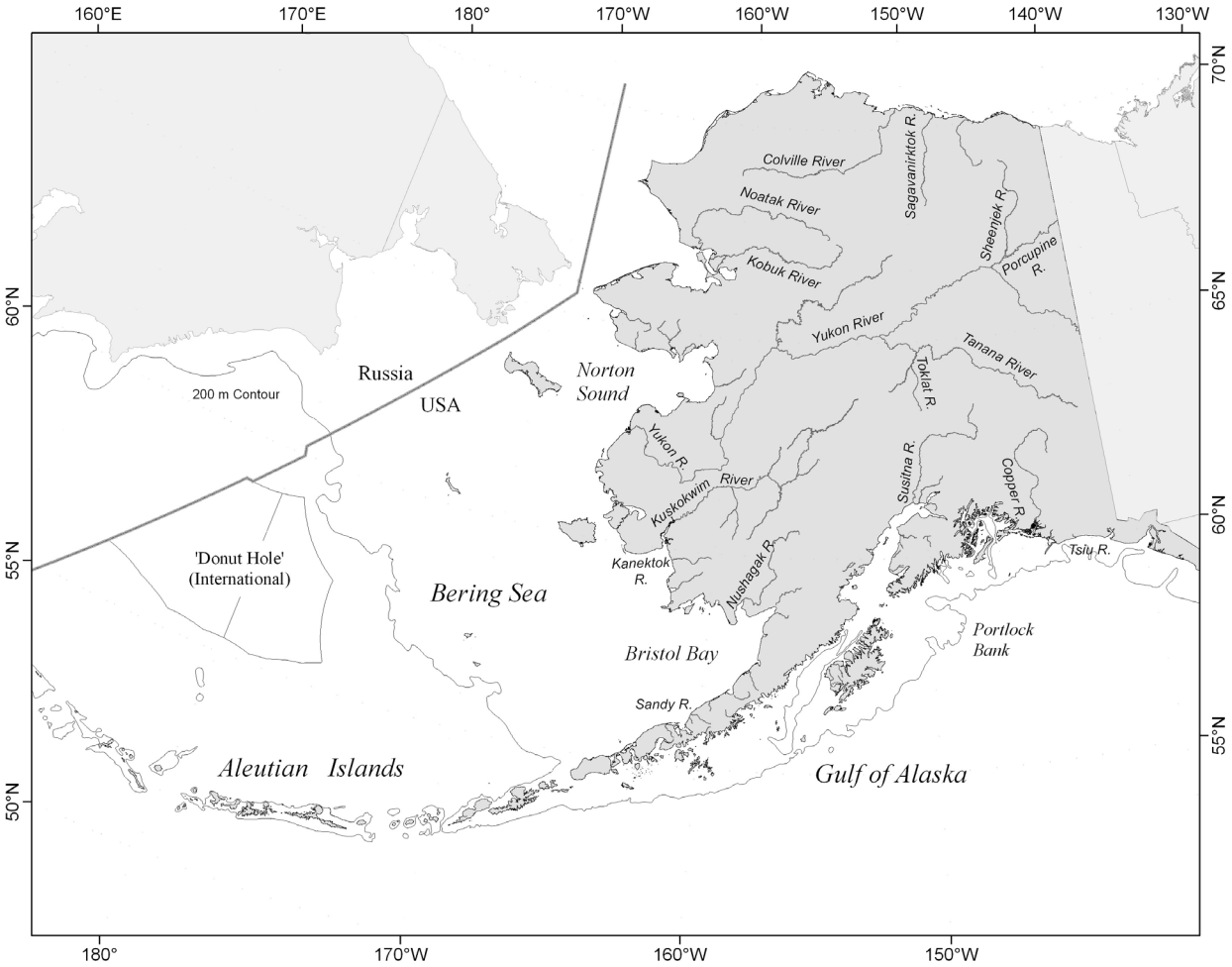

Figure 1.–Major fishing regions and important river systems of Alaska.

The American Fisheries Act of 1998 concerned the Bering Sea–Aleutian Islands pollock fishery and specified vessel ownership provisions, pollock allocations, vessel retirement and buyout, co-ops, and so-called sideboards to protect other fisheries from unintended effects of the American Fisheries Act. The American Fisheries Act specified an allocation of the directed pollock fishery annual *Total Allowable Catch*: 10% of *Total Allowable Catch* went to the Community Development Quota Program. The remaining *Total Allowable Catch* was divided among inshore, offshore, and mothership components at 50%, 40%, and 10%, respectively. The American

Fisheries Act limited the number of vessels targeting Bering Sea–Aleutian Islands pollock and allowed the formation of cooperatives.

Current Federal regulations divide this series of fisheries into various overlapping and non-overlapping categories and groupings for different purposes. Many of the relevant rules that govern these fisheries can be found in Federal Fisheries Regulations, 50 CFR Part 679: Fisheries of the Exclusive Economic Zone Off Alaska (http://www.fakr.noaa.gov/regs/default.htm) (hereafter, unless otherwise specified, citations with a § symbol will refer to specific sections of these Federal regulations). These commercial groundfish fisheries are sometimes described by gear used, either trawls, pots, or longlines (Woodby et al. 2005). The commercial fishery is often described in terms of *sectors*, and within the Bering Sea–Aleutian Islands fishery, these sectors have traditionally been described as (1) the inshore, (2) the offshore, and (3) the mothership sectors of the industry. These terms are repeatedly used in Federal regulations (§679.2). Recently, the Community Development Quota fisheries have been described as the fourth sector. The inshore sector is comprised of processing facilities that are located on land as well as two stationary floating processors, the *Arctic Enterprise* and the *Northern Victor*. The inshore sector, also commonly referred to as the shoreside sector, relies on catcher vessels that are separated in regulation by their length overall into two categories, and these catcher vessels reported the highest Chinook-to-pollock catch ratios in 2007 (Table 1).

The State of Alaska manages groundfish fisheries within state territorial waters (within 0–3 nm), while most fisheries within the U.S. Exclusive Economic Zone (3–200 nm offshore) fall under Federal management. The management of fisheries in Federal waters is controlled through plans approved by the North Pacific Fishery Management Council.

Table 1.–Catch and bycatch statistics from the 2007 Bering Sea pollock fishery, by vessel category.

Vessel Category	Number of vessels	Pollock catch (thousands of metric tons)	Chinook bycatch (thousands of fish)	Chinook ratio (number per metric ton of pollock)
Catcher-processor	16	608	35	0.06
Mothership	3	141	9	0.06
Shoreside	83	573	77	0.13
Total	102	1,322	121	

Source: http://www.fakr.noaa.gov/sustainablefisheries/bycatch/salmon/chinook/bycatchbyseason_sector02_07.xls , last updated 6/2/2008

The Bering Sea–Aleutian Islands pollock fishery is opened during two seasons of the year, called the *A Season* and the *B Season*. The A Season runs from January 20th to June 10th. During this season, pollock are in spawning condition, and the fish are moving into spawning aggregations. Because the fish roe is valuable in Japanese markets, the catch is more valuable per unit weight during the A Season. The B Season runs from June 10th to November 1st. During this season, the primary product is fish flesh.

THE BYCATCH

The notion that certain species were to be prohibited in the groundfish fisheries dates back to the 1940s and 1950s (Witherell and Pautzke 1997), and the original management plan for the Bering Sea–Aleutian Islands groundfish fisheries prohibited the retention of salmon. Recognizing that

some level of incidental harvest of prohibited species is necessary to efficiently harvest ground fish, in 1982 the level of Chinook salmon bycatch was capped at 55,250 for foreign-owned trawl vessels fishing in the Bering Sea–Aleutian Islands fisheries, with the available bycatch apportioned to various fishing nations (Witherell and Woodby 2005). In later years, after the groundfish fishery was fully Americanized, the incidental harvest of Pacific salmon in the combined Bering Sea–Aleutian Islands and Gulf of Alaska groundfish fisheries averaged 38,000 Chinook salmon and 69,000 "other" (mostly chum) salmon. Between 1990 and 2001, on average, slightly more than a third of the Chinook salmon bycatch was taken in the Gulf of Alaska (Witherell et al. 2002).

In 2000, a new system of regulation was put in place to reduce the bycatch of Chinook salmon in the Bering Sea–Aleutian Islands trawl fisheries. The allowable Chinook salmon bycatch—described as a "limit" by Witherell et al. (2002)—was to decrease from 48,000 to 29,000 over a 4-year period. The 2000 plan created areas called *Chinook salmon savings areas* that would be closed to pollock fishing in years with high bycatch, and the plan also created more restrictive closing dates (Witherell et al. 2002). The timing of the closures depended on when the limit was reached. If the limit was reached before April 15th, the Chinook salmon savings areas were to close immediately, stay closed through April 15th, and then close again from September 1st to December 31st. If the limit was reached after April 15th, but before September 1st, the Chinook salmon savings areas would close on September 1st through the rest of the year. If the limit was reached after September 1st, the area was to close immediately and stay closed through the year.

Since these changes in 2000, the estimated overall level of Chinook salmon bycatch has strongly trended upward (Anonymous 2008), although the bycatch dropped dramatically in 2008. The bycatch estimate has either exceeded or else greatly exceeded this level of 29,000 every year since 2001, with large catches in both the A Season and B Season (Table 2). Chinook salmon bycatches tended to be higher in the A Season, although in some years (e.g., 2006) the bycatch level was similar in both seasons. In 2004 and 2005, the bycatch was higher in the B Season. In 2003, the Chinook salmon bycatch estimate reached 54,000 for the entire Bering Sea–Aleutian Islands groundfish fisheries (not just the directed pollock fishery), with this level approaching 46,000 in the directed pollock fishery in this same year. This level of bycatch in 2003 triggered closures of a Chinook salmon savings area for the first time. The bycatch levels further increased in 2004, resulting in time and area closures again (Witherell and Woodby 2005).

Similarly, the estimated bycatch of salmon classified as "other" (mostly chum salmon) has strongly trended upward since 2000, with the 2005 peak level of 705,000 salmon in the directed Bering Sea–Aleutian Islands pollock fishery bycatch equaling more than 10 times the 2000 level of 59,000. The bycatch level has declined since 2005: the 2006 estimate was 310,000 and the 2007 estimate was 94,000 in the directed pollock fishery—still quite high by historical standards (Table 2). For those species classified as "other" salmon, the bycatch is higher in the B Season by a factor that is typically between 10 and 100 (Table 2). In the early 1990s, the inshore catch of chum salmon in Western Alaskan salmon fisheries sharply declined due to decreased abundance, and this coincided with a sharp spike in the bycatch of chum salmon in the directed pollock fisheries. Not surprisingly, in the early 1990s there was a great deal of interest in the bycatch of the salmon classified as "other" (non-Chinook, mostly chum salmon), but currently most of the concern is with the Chinook salmon bycatch.

Table 2.–Estimated salmon bycatch in Bering Sea–Aleutian Islands directed pollock fisheries in thousands of fish. The estimates include catches in the Community Development Quota sector of the fishery.

	Chinook Salmon			"Other" (mostly chum) salmon		
Year	Annual	A Season	B Season	Annual	A Season	B Season
2000	5	3	2	59	<1	58
2001	33	18	15	57	2	55
2002	34	22	13	81	1	79
2003	46	33	13	189	4	185
2004	52	23	29	440	<1	440
2005	67	27	40	705	<1	704
2006	83	58	24	310	1	308
2007	121	69	52	94	9	85

Source: Salmon bycatch was described as "salmon mortality" in the original source. National Marine Fisheries Service (http://www.fakr.noaa.gov/sustainablefisheries/inseason/chinook_salmon_mortality.pdf), updated 3/19/2009.

As an alternative to the time and area closures to conserve Chinook salmon, representatives of the pollock fleets proposed a voluntary system of rolling area closures based on actual measured bycatch rates. This system is sometimes called the inter-cooperative agreement (or ICA) or the *rolling hot spots* system. Vessels that participate, and agree to rolling area closures, have been exempted from closures in the Chinook salmon savings areas. In 2006, the catcher vessels contracted with an independent firm, Sea State Inc., to coordinate this effort, and vessel owners participating in this program agreed to let the National Marine Fisheries Service release confidential information to Sea State Inc. This firm currently monitors individual vessel bycatch rates, and then reports aggregate rates to the vessels in real time. Sea State does not receive information for vessels that have not provided clearance. Sea State also published a "Dirty 20" list of vessels with highest salmon bycatch rates. In April 2009, the North Pacific Fishery Management Council imposed an annual bycatch limit of 60,000 Chinook salmon for the Bering Sea pollock trawl fishery and adopted a Chinook salmon bycatch management program with incentives for reducing salmon bycatch at all levels of salmon abundance and salmon encounters.

THE OBSERVER PROGRAM

Although U.S. observers had monitored foreign vessels in Alaskan groundfish fisheries as early as 1973, a joint industry and National Marine Fisheries Service pilot program to monitor the later U.S. vessels that replaced the foreign fleet first began to collect data needed to manage the fisheries in 1989. By 1990, the North Pacific Groundfish Observer Program (hereafter called the "Observer Program") was established in its current form under the National Marine Fisheries Service (Barns et al. 2005). The Observer Program is headquartered at the Fisheries Monitoring and Analysis Division of the Alaska Fisheries Science Center in Seattle. The program provides essential data on operations of the domestic commercial fisheries off Alaska that is used in fisheries management, stock assessment, and scientific research: it monitors prohibited species catch, collects biological data and samples for stock assessment, documents fishery interactions with protected species, collects samples and data for marine ecosystem research, and conducts fishery compliance monitoring. In particular, the Observer Program provides the pollock catch and salmon bycatch information that is the focus of this study. The National Marine Fisheries Service certifies contractors that provide the observers, and then trains and certifies the observers

as well. An extensive manual is produced annually (2008 Observer Sampling Manual is downloadable as a 465 page PDF document from http://www.afsc.noaa.gov/ FMA/ document.htm.) that details the observers' duties, data to be collected, sampling methods to be used, forms to be completed, as well as listing fishery regulations. This manual series documents the improvements that have been made to the sampling methods in response to user's needs, recommendations from independent reviewers (Vϕlstad et al. 1997 and the Marine Resources Assessment Group Americas, Inc., or MRAG 2000, 2003), and recommendations from in-house staff. Unless otherwise cited, most of the material in this section can be found in this Observer Manual.

At this time, the National Marine Fisheries Service has approved five contractors (Table 3). The contractors provide over 400 observers for the program. These observers expend more than 35,000 observer-days monitoring vessels and processing plants in the North Pacific (Cahalan and Ferdinand 2008). The fishing industry currently pays for the observer services. Vessels and plants that are required to have observer coverage must arrange for an observer with a certified provider, and pay the provider for the observer services. National Marine Fisheries Service employees that work in the program are stationed in Seattle, Anchorage, Dutch Harbor, and Kodiak. The National Marine Fisheries Service pays the combined costs for administering the program, the training, briefing, and debriefing of observers, and management of observer data (Barns et al. 2005).

Table 3.–Contractors with permits from the National Marine Fisheries Service to provide observers for the North Pacific Groundfish Observer Program.

Alaskan Observers, Inc. 130 Nickerson, Suite 206 Seattle, WA 98109 Phone: (206) 283-7310 www.alaskanobservers.com	Saltwater, Inc. 733 N. Street Anchorage, AK 99501 Phone: (907) 276-3241 www.saltwaterinc.com
NWO, Inc. P.O. Box 624 Edmonds, WA 98020 Phone: (425) 673-6445 www.nwoinc.com	TechSea International Corp. 12510 33rd Avenue NE Suite L-103 Seattle, WA 98125 Phone: (206) 285-1408 www.TechSea.com
MRAG Americas, Inc. 1810 Shadetree Circle Anchorage, AK 99502 Phone: (907) 677-8772 www.mragamericas.com	

Source: Groundfish Observers Manual.

Processors, fishing vessels, and their coverage and information through the Observer Program

Observers gather information about the pollock fishery and its bycatch at the processing plants and on the fishing vessels. In the course of their duties the information they gather is recorded on a number of different forms (see Appendix A for a partial listing of these forms and brief description of the information gathered). The finest resolution of information is the catch in

individual trawl hauls by the participating vessels. In the descriptions of observer activities that follow, we focus on the resolution possible for fishery operations that are available for the observers to obtain the data needed to estimate bycatch and its geographic origins. The sampling situations that the observers encounter can be divided into two categories: (1) the catch of each trawl haul by a fishing vessel can be sampled separately from that in the other hauls, and (2) only the combined catches of all hauls during a vessel trip can be sampled. The first category further divides into two subcategories: (1) the unsorted catch of each haul can be sampled, and (2) the catch may have been sorted before a haul can be sampled. Sorting the prohibited salmon bycatch from the target pollock catch will always be done at some place and time after the haul is brought on board and when the catch is processed at the plant.

Motherships are at-sea processors that rely on catcher vessels to harvest and deliver fish to them. The catcher vessels that deliver to motherships do not require observers because their trawl codend is passed with the unsorted catch to the mothership. Three American Fisheries Act motherships participate in the pollock fishery, and each must have at least two observers aboard for each day the vessel is used to process or take deliveries of groundfish. Each observer covers a 12-hour shift, and all hauls delivered from the catchers are observed and sampled unless the observer is incapacitated. The motherships must provide a motion compensated flow scale on which all catch in each haul is to be weighed. Also, an observer sampling station is required that includes a motion compensated platform scale and other tools to assist in sampling the catch. With a platform scale, the amount of fish that can be sampled is limited by the observer's capacity to move the fish, allowing about 300 kg per sampled haul.

Catcher-processor vessels catch fish, process them at sea, and freeze their product. American Fisheries Act catcher processors have the same observer and monitoring requirements as the motherships. Therefore, the individual hauls can be sampled before sorting.

Catcher vessels are smaller than catcher-processor vessels, and capture and deliver their catch to catcher-processors, motherships, and floating and inshore processors for processing. A delivery may involve simply passing the unsorted catch in the trawl codend to the mothership processor, or the catch may be removed from the trawl, stored in holding tanks, and then pumped from the holding tanks to the shore-based or floating processor.

Processing plants are of two kinds: floating (or at-sea) and shore-based. Observer coverage of processing plants during the pollock open season is based on the amount of groundfish delivered to them each month. Plants that process 1,000 metric tons or more of groundfish per month must have an observer present each day they process fish. Plants that process between 500 metric tons and 1,000 metric tons each month must have an observer present for 30% of those days.

Some catcher vessels deliver their catch to floating processors. Vessels that routinely pump fish from delivering catcher vessels are considered to be floating processors. The fish delivered to these vessels have the potential to be sorted at sea before delivery, and in this regard the delivery to the floating processor is like that to a shore-based processing plant. Any sorting that was done before delivery would be unknown unless an observer was on board the catcher vessel.

Some catcher vessels deliver their catch to shore-based processors. The catches from hauls of the trip are not distinguishable, and sorting that may have occurred would be unknown unless an observer was on board the catcher vessel. On days for which shore-based processors are required to have an observer present, they must provide an observer for each 12-hour shift of each day, during which the processor takes delivery of, or processes, groundfish from a vessel fishing

pollock. The shore-based processors must also provide a location from which the observer can see all sorting and weighing of fish simultaneously. The observer meets every vessel arriving at a plant. Plant personnel weigh the catch as it is offloaded using scales certified by the State of Alaska. The observer checks that these weights are recorded accurately, and that the entire delivery is weighed. The observer at a plant also collects and records Offload Form (see Appendix A) data from captains, logbooks, or vessel observers, when present; assists vessel observers during offloading; and collects lengths and age structures. Here is where information from vessels with or without observers aboard is obtained. At minimum, delivery information on the Alaska Department of Fish and Game fish ticket is available if additional information from the other sources is not forthcoming. An after-scale area is where plant workers set aside bycatch that made it past the sorters and into the factory. This location is where some salmon bycatch may be discovered. The plant observer is asked to record the species, length, weight, and sex of prohibited species, and to give this information together with the rest of the delivery information to the vessel observer, if one was aboard.

Bering Sea–Aleutian Islands salmon regulations require operators of vessels carrying an observer to retain all salmon separate for each haul, if sorted from the catch, until the observer can count and sample them. Operators of vessels not carrying observers onboard, or whose fishing operations do not allow for sorting of salmon from the catch, must store the salmon for delivery to the processor. If an observer is at the receiving plant, that observer will count the delivered salmon bycatch. Therefore, the count of Chinook salmon bycatch in each haul from a trip with an onboard observer should, in principle, be known and available at the time of statistical analysis. In principle, for trips without an observer aboard, an observer at the receiving plant should have collected the individual trip or delivery bycatch statistics, if an observer was present (see § 679.21).

Sampling methods on trawlers

During an observer trip, if less than 100% of hauls are sampled, the hauls to be sampled are selected from the total hauls during the trip with a protocol that introduces chance. Then the observer assesses the best method for sampling a haul and establishes the sampling frame to be used, either temporal (units defined by time), spatial (units defined by location or space), or a combined temporal-spatial frame. Finally the observer selects at least three species-composition samples (collections or clusters of fish) at random from the potential units of the frame. The samples are weighed, and for each salmon species present, the number, sex, and length of each fish are recorded. Scales are also taken from up to 20 individuals of each salmon species per trip, and the length, weight, and sex of each is recorded.

The Fisheries Monitoring and Analysis Division databases as of January 2008

The data from observers are stored on two databases under the Information and Monitoring Technologies Program. First, observers enter and transmit their data to the at-sea database. Second, the repository database at the Alaska Fisheries Science Center in Seattle stores the data permanently. Generally, observers now record only what they see or measure. Previously, computed, interpolated, or imputed values were entered into the databases by observers, and these values could not be distinguished from actual measured data points. The automation of most computation within the at-sea and Alaska Fisheries Science Center databases increased flexibility and reduced errors. The database now uniquely identifies observed trips and uniquely identifies collections (such as the salmon tissues) taken from observed hauls. Trip-specific

imputation methods can be used in cases where not all hauls are sampled for a trip. Design-based estimators of variance are possible because of multiple species composition collections from each sampled haul.

The structure of the redesigned database mirrors observer sample collection. Trips are linked to observer cruise number and vessel, hauls are linked to trips, and sample collections are linked to hauls. Other data collected, such as scales and tissue specimens, marine mammal information, and fish lengths are recorded in relation to where they were obtained. For example, the scales and length data can be related to a sample, a haul, or an offload. Observers are now asked to take at least three sample collections from groups of fish from each fishing event, unless it is physically or logistically impossible to do so. The National Marine Fisheries Service has now moved to a preference for selecting sample collections using systematic spacing throughout the hauls, beginning at a random starting position to capture any stratification or sorting (Cahalan and Ferdinand 2008). To be clear, the systematic sampling is of clusters of fish of mixed species—not individual salmon.

SPECIFIC SAMPLING TERMS AND ESTIMATION ISSUES

In designing an effective sampling study, our path of logic begins with a clear and achievable goal, an assessment of the resources available, and a plan for using those resources to reach the goal. A scientific sampling study should begin with an idea of a *statistical population* made up of *sampling units*, each of which has an associated value. For example, the statistical population could be the vessel deliveries of fish to processors in the pollock fishery during the 2008 B Season, and the associated value of the sampling units (i.e., the deliveries) might be a Chinook salmon bycatch rate or ratio (e.g., the number of Chinook salmon per metric ton of pollock). Next, the scientific sampling study requires an idea of *population parameters* (the numerical summaries for the entire population), such as the actual mean and variance of the Chinook salmon bycatch ratio for deliveries during the B Season. Next, the scientific sampling study requires an idea of a *sample*, or a subset of the sampling units selected from their statistical population and for which the associated values are measured. Last, the scientific sampling study requires an idea of *parameter estimates* (numerical summaries or functions of the sample data), which are approximations of the unknown population parameters based on the measurements in the sample. An organized and disciplined program for the collection of the basic sample data is essential, and a logical way to connect the sample data to the population parameters and the research questions has to be available. The *statistical design* is the plan for selecting a representative sample of sampling units from the population by which to infer probable values of the population parameters and assess their uncertainty (Thompson 2002). A foundational principle of scientific sampling is that the unsampled values may only be considered similar to the sampled values if assurance is available that the sample is somehow representative of whole population.

The gold-standard selection technique to assure the sampling units are representative of the statistical population is *random sampling*, for reasons that have been accepted for almost 100 years (Fisher 1925; Cochran et al. 1954; Cochran 1977; Thompson 2002; and many others). This notion of a *random sample* has been expanded and generalized into what is sometimes called *designed-based sampling theory* (Thompson 2002), which includes the topics of stratified random sampling, cluster sampling, unequal-probability sampling, adaptive sampling, and many others. The unifying feature of all of these approaches is the principle that by selecting each

sampling unit for inclusion in the sample with a protocol of randomization, features of the original population are preserved in the sample, including the variation among sampling units. The variation among units within the sample is used to study the effectiveness of the sampling procedure, and a measure of that variation can be re-expressed in the form of confidence intervals, standard errors, and so on. In everyday speech the term *random* can mean "without method or conscious decision." However, the term *random sample* has a very specific—and quite different—technical meaning: for statistical sampling theory in general and the genetic analysis described later in particular, the term *simple random sample* means that a very definite protocol was used to include sampling units in the sample by chance in such a way that every sampling unit in the population had the same probability of entering the sample.

With random sampling that includes a large enough sample size, in some precise mathematical sense, the sample will be assured to be representative of the population as a whole. Moreover, with a protocol of randomization, in some precise mathematical sense, the parameter estimates will, on average, predictably get closer to the true population parameters as the sample size gets bigger. For example, if the total B Season Chinook salmon bycatch in the Bering Sea–Aleutian Islands pollock fishery were the statistical population, and stock origin of a fish was its associated value, a simple random sample of several hundred individuals drawn from the tens of thousands of individuals would be guaranteed to have a stock composition very close to that of the total B Season bycatch. For operational reasons, systematic samples (see below) can sometimes be assumed to be equivalent or even superior to random sampling (Cochran 1977). However, it is strictly inappropriate to assume that any sample of convenience is in fact a random sample (Cochran et al. 1954; and many others).

Another branch of scientific sampling, called *model-based sampling*, rests on the idea that the statistical values in the population may be thought of as random variables coming from some probability distribution with unknown parameters. This approach can be expanded into techniques that rely on some relation between variables (some of which can be measured) and the underlying parameters of the probability distribution (which cannot) (Thompson 2002). An instance of model-based sampling would be filling in for Chinook salmon bycatch in unsampled vessel trips with estimates from a linear model relating Chinook salmon bycatch to pollock catch. Model-based sampling can often be the best way to use the data that is actually available.

Often the model-based sampling does not require the assumption that the sample was selected as a random sample. However, this approach will usually depend on a series of complex and important assumptions that must be examined before the statistical estimates can be viewed as realistic or appropriate. As Thompson (2002) points out, when the results of sampling studies must be accepted by people of conflicting viewpoints and economic interests, the assumptions underlying model-based sampling may not be acceptable to all parties. He points out that introducing an actual randomization process in the selection of sampling units provides a powerful safeguard that the resulting parameter estimates fairly represent the actual unknown parameters of the problem. This protection is an important consideration when allocating resources, such as fish (Thompson 2002, Section 10.1).

Currently, in the Bering Sea groundfish fisheries the process of sampling occurs at several different levels and the word *sample* is often confusing when applied to individual fish, tissue specimens, vessels with observer coverage, collections of individual fish or vessels, or many other things. Below, when we use the word as a noun we will attempt to use the term *specimen*

or *sampling units* if we mean individuals within a sample or a population and the word *sample* if we mean the entire collection of specimens that will be used to estimate population parameters.

We will use the term *sampling domain* to mean a collection of sampling units that are identified in some way (maybe because of their relationship in time and space) because we are interested in separate parameter estimates for that particular collection. For example, if we think of the catcher vessel fleet's deliveries as the statistical population, then we may want to divide this population into time and area domains, such as deliveries during B Season, or deliveries from particular management areas. A critical reading of "Genetics-Based Information on Stock Origins of the Bycatch" will show that we will need to produce a collection of scales and tissue specimens that can be thought of as a random sample drawn from all Chinook salmon caught in the same time and area domain, irrespective of haul, delivery, or vessel.

A recurring theme throughout the rest of the document will be that sampling will take place at many different levels. For example, an onboard observer's job may require thinking of one specific haul as a statistical population made up of large number of potential cluster samples (sometimes called basket samples) of multi-species assemblages. However, that haul will just be one of several hauls for the entire trip. The hauls may be thought of as sampling units if not all hauls can be examined because of workload. Because total Chinook salmon bycatch will be recorded for the delivery, which may be made up of several hauls, for the purpose of estimating bycatch in this catcher vessel's size class, the deliveries can be viewed as the sampling units. Managers may require grouping some of these hauls into a statistical domain for additional study, say, to examine the bycatch level in an area that might be closed for salmon conservation. The ability to match the sampling units across levels of sampling will be essential when estimating the bycatch by its regional geographic origins for sampling domains. On one hand, deliveries or hauls are the sampling units when the statistical population is the total bycatch in a domain, while on the other hand, individual salmon in those hauls are the sampling units for genetic decoding of bycatch origins in the domain. For the purpose of genetic stock identification, the individual Chinook salmon are the sampling units and the statistical population is the entire collection of individual salmon within the domain. However, this statistical population will be physically dispersed among vessels or processing operations at the point when observers will be able to collect specimens.

THE CURRENT APPROACH TO ESTIMATING A BYCATCH TOTAL

To develop estimates of bycatch by its regional geographic origins, first we must obtain a reliable estimate of total Chinook salmon bycatch. Using a design-based approach, we might think of vessel deliveries (for a vessel category with incomplete information) as the sampling units, and we might think of the bycatch in each delivery as the value associated with the sampling unit. Then, if bycatch was measured only for some deliveries, we might consider these deliveries to be a random sample from the larger population of all deliveries, both those for which bycatch was measured and those for which bycatch was unknown. If we are convinced that the sample resembles a random sample, then the total bycatch for the vessel category could be estimated as the total number of deliveries times the average Chinook salmon per delivery from the sample. Alternatively, from the model-based point of view we might group all deliveries into categories (based on time and area fished), and assume for each category that each delivery's ratio of Chinook salmon to metric tons of pollock is a random variable drawn from a category-specific probability distribution with fixed parameters. The bycatch of each delivery in

each category can be estimated using these distributions. First, we could produce an estimate of the mean Chinook salmon ratio of each category's probability distribution. Then we could treat this estimate as a regression parameter, and we could estimate the bycatch of each delivery by multiplying the category regression parameter by the measured weight of the pollock catch for each delivery in the category. This would be like treating the weight of pollock for each delivery as an independent variable in a regression model. For the model-based approach, the bycatch estimate would be found by adding these products for the deliveries with unknown bycatch in all categories to the bycatch in deliveries for which it was measured.

For the last few years, the National Marine Fisheries Service has apparently estimated Chinook salmon bycatch using a style that seems more like the model-based than design-based approach. The method is used in their computer-based Catch Accounting System (Anonymous 2008). They developed a series of estimates of Chinook-to-groundfish ratios. These ratios were developed using only information collected from trips or deliveries with an observer onboard. The National Marine Fishery Service refers to these ratios from deliveries with onboard observers with a series of non-descriptive database codes, such as *Precedence 20*, *Precedence 25*, and so on. For example, the *Precedence 50* refers to a bycatch rate calculated by using information that is aggregated with the following characteristics: "Vessel specific, date trip started, fishing gear, federal reporting area[1]." The National Marine Fishery Service has not documented how well these bycatch rates perform when expanded into estimates of bycatch for specific subsets of the fishery. We do believe that sound estimation methods for the observer program that were developed by Miller (2005) are not currently used. Therefore, in Appendix B we will review parts of Miller's PhD thesis, as it applies to salmon bycatch in the pollock trawl fishery. Miller's 419-page compendium of the sampling theory and estimation formulas is broader in coverage than needed here. In Appendix B we have attempted a simplification of his notation, and we have restricted attention to 23 key equations.

ALTERNATE INFORMATION ABOUT SALMON BYCATCH

A different way to obtain an estimate of total Chinook salmon bycatch is to simply sum the observers' counts of Chinook salmon for every delivery (for catcher vessels) or haul (for motherships and catcher-processors). Regulations (§ 679.50) are in place requiring observers on board during most mothership and catcher-processor operations, and two observers are required during most Community Development Quota fisheries. Similarly, observers are carried on all trips for the larger catcher vessels. So the estimate of bycatch is essentially a census for these vessel categories. For midsized catcher vessels, 70% of the trips do not have onboard observers. However, regulations require that "…the operator of a vessel and the manager of a shoreside processor must not discard any salmon or transfer or process any salmon under the [prohibited species donation program], if the salmon were taken incidentally to a directed fishery for BSAI groundfish by vessels using trawl gear **until the number of salmon has been determined by an observer** and the collection of any scientific data or biological samples from the salmon has been completed," (§ 679.21).

In other words, if this regulation is enforced, then the estimation of bycatch is—in principle—essentially an accounting problem and not one requiring probability-based statistical sampling for any of the vessel categories, including the midsized catcher vessels. Perhaps it is unrealistic

[1] From an unpublished National Marine Fisheries Service handout from the February 2008 North Pacific Fishery Management Council meeting.

to think that this regulation could be enforced, given practical constraints on the whole Observer Program. However, ideally, individual bycatch would be attributable to every delivery. In statistical terms the parameter estimates should be obtainable this way with little or no sampling error (error due to the fact that only a subset of the statistical population was used to make the estimate), and any error would be the result of non-sampling errors (transcription errors, miscoding, database management problems, intentional misrepresentations, etc.).

Miller (2005) developed estimates of Chinook salmon bycatch from 1993 to 2003 that were considerably different from those of the National Marine Fisheries Service. In every case Miller's estimates are substantially larger. For example, Miller estimated the Bering Sea bycatch in 2003 for longline and trawl gear to be more than 79,000 with standard error of about 3,000 (see Miller's Table 6.23), while the all-gear Bering Sea and Aleutian areas estimate from the National Marine Fisheries Service was approximately 55,000 with unstated precision. These discrepancies point to the need to develop written procedures for the process that the National Marine Fisheries Service uses to produce bycatch estimates. This inconsistency also points to the need to develop written procedures for the assessment of the quality of the data and resulting estimates, both in terms of sampling errors and non-sampling errors.

IMPORTANT TESTS OF ASSUMPTIONS

An unstated assumption underpinning the National Marine Fisheries Service's estimates of bycatch is that the rate of salmon bycatch is essentially the same for trips with and without onboard observers. The *observer effect* is the term for a hypothetical difference in these rates that might be caused by vessel operators of midsize catcher vessels altering their fishing operations for the 30% of trips with onboard observers. This observer effect, if it occurs, will obviously lead to an underestimate of salmon bycatch. That will be true using the methods of Miller (2005) or presumably the methods that the National Marine Fisheries Service has been using the last several years, and this underestimate will occur whether using design-based methods or model-based methods. However, this observer effect should be detectable through an appropriate statistical test, especially if the practice is widespread. Such a test would depend on enforcement of the regulations requiring a determination of the salmon bycatch on trips without onboard observers. Yet enforcement of this regulation (§ 679.21) would then also allow bycatch estimates to be viewed as a census—that is, without sampling error.

Put another way, two critical assumptions are necessary to view the cumulative Chinook salmon bycatch as a census: midsized vessels must obey the regulations regarding salmon bycatch (i.e., the salmon bycatch must actually be delivered to the processing plant during those 70% of a vessel's fishing days for which no observer is on board), and observers must be clear that they expect the regulation to be followed, report violations, and ensure that the delivered salmon are counted and the counts are recorded to the appropriate databases. Assuming these counts are made, if there is an intentional misrepresentation or intentional discarding of salmon before observers count the bycatch, that condition should be detectable through an appropriate statistical test, especially if this practice is widespread. Because there could easily be legitimate reasons for differences in bycatch rates among vessels fishing in the same general area, these tests are constructed with that consideration in mind.

Although many possible statistical tests might be appropriate, a simple starting point would be to average the Chinook-to-groundfish ratios, on a vessel-specific basis, for those trips with observers and those trips without observers. For each quarter in 2007, 57 vessels required

observer coverage on 30% of their cruises (Table 1), so we could assemble approximately 57 of these paired averages per quarter. If a vessel is equally likely to have a slightly higher ratio just due to chance alone when an observer is on board as when an observer is not on board, then on average one-half (28.5) of the 57 paired averages should be higher when an observer is on board, and one-half should be smaller. Moreover, intentional underreporting on unobserved trips and the occurrence of altered fishing operations when observers are on board would cause the ratio to change between trips with or without an observer: vessels would have a higher average ratio during trips with onboard observers than without. The binomial probability distribution can be used to evaluate the probability that any number of vessels among the 57 would have a ratio as high or higher if due only to chance.

When there are two possible outcomes with fixed probabilities of occurrence at each of a specific number of random trials, then the binomial probability distribution can be used to evaluate the probability for getting a specified total number of either outcome. For example, the binomial probability distribution provides the probability of getting a particular number of heads when a fair coin is flipped 57 times: the total number of observed heads should fall between 23 and 34 over 80% of the time. So, fleet wide, if more than 34 vessels returned with a higher bycatch ratio when observers were on board, we could conclude that there was very weak evidence that salmon had been discarded when observers were not on board. Similarly, if 23 or less returned with a lower bycatch ratio when observers were on board, we could conclude that there was very weak evidence that fishing operations had changed for trips with observers, because being outside the range of 23 to 34 should happen less than 20% of the time due to chance alone. However, if 20 or less, or 37 or more vessels had a higher bycatch rate with observers on board, we could conclude that there was moderate evidence of one of these problems (the count of vessels with such unusual ratios should happen less than 5% of the time due to chance alone), and if 16 or less, or 44 or more vessels had a higher bycatch ratio with observers on board, we could conclude that there was very strong evidence of one of these problems (the count of vessels with such unusual ratios should happen less than 0.2% of the time due to chance alone). Of course, very similar ideas could be used to track an individual vessel's performance over time or to test for only one hypothesis at a time (see Edgington 1980 for a more general discussion of these kinds of tests). The fact that a check for at-sea dumping is so obvious and readily available should strongly deter the illegal practice, if it occurs at all, given the additional publicity in an already visible issue. However, these kinds of binomial tests are not possible without a census approach to accounting for bycatch for each delivery, irrespective of whether an observer was on board or not.

IMPROVED BYCATCH ESTIMATES WITH INCOMPLETE INFORMATION

As we previously noted, managers and members of the public are interested in salmon bycatch by the trawl fishery for pollock at several levels. First, an overall accounting of bycatch removals by its regional geographic origins is needed, possibly stratified by vessel categories. Second, for possible inseason management actions to effect reductions in bycatch, the focus will be time and area domains. The fishery during a season may be concentrated in areas where unfortunately the salmon are abundant and are caught at high rates. The fishery may need to move until the salmon abundance declines in such places if the bycatch is to be limited. Moreover, there may be certain times and areas where stocks of particular concern were more vulnerable. The particular times and areas of high salmon abundance could change among years so the sampling program to estimate salmon bycatch has to be flexible in order to adapt to these changes. Third, and last, at

either intra-seasonal or seasonal levels, information regarding the population sources of the salmon bycatch will be invaluable in formulating management decisions.

The cumulative Chinook salmon bycatch, by day of season, equals the sum of the bycatch for all trawl hauls by the fleet to that time. A reading of the Observer's Manual and Bering Sea–Aleutian Islands salmon regulations indicates that the haul counts should be available each day for each large vessel 125′ or greater that sorts its catch at sea. A *haul count* is the complete census of salmon individuals in a haul. The haul counts also should be available for the 30% of days that an observer is on board a midsized (60′ to 125′) vessel that sorts its catch at sea. If these vessels of either size class do not sort at sea, or if no observer is on board a midsized vessel (an observer is always on board a large vessel) under the salmon regulations, the salmon bycatch must be stored until delivery to a processing plant. Therefore, the trip counts of salmon should be available if delivery is to a plant with an observer. Plants that process 1,000 metric tons or more of groundfish per month must have an observer present each day that they process fish. Plants that process 500 to 1,000 metric tons per month must have an observer present for 30% of those days. Rare combinations of no sorting at sea, no observer on board a midsized vessel, or no observer at a plant would allow for some bycatch being missed in the census. However, according to the most recent draft of the Environmental Impact Statement for the Bering Sea–Aleutian Islands Chinook salmon bycatch management (Anonymous 2008), "For the pollock fishery, the vast majority of tows are observed either directly at sea or at offloading locations aboard motherships or at shore-based processing plants." Therefore, if deliveries to plants without observers are assumed to be negligible and small vessels <60′ are not present in the fishery, the trip counts of salmon bycatch are in principle known for all vessels throughout the season. A trip count is the complete census of the salmon individuals caught between departure and return to a processing plant.

Despite the essentially complete census of bycatch that seems to us available, others have indicated that the census is less complete than regulations suggest it should be. The National Marine Fisheries Service Regional Office in Juneau computes estimates of the Chinook salmon bycatch using their computerized Catch Accounting System (Anonymous 2008). The Observer Program provides input data for the Catch Accounting System about pollock catch and Chinook salmon bycatch during observed vessel trips (from the Observer Manual—*op. cit.*). These data include pollock catch and Chinook salmon bycatch estimates per sampled haul (which are expansions of observations in species composition samples taken from each haul to account for the unobserved fraction of the haul) as well as corresponding estimates for any unsampled hauls during the trip. As we understand it, the Catch Accounting System has an algorithm by which to estimate Chinook salmon bycatch in hauls by vessels without an observer. The algorithm computes a surrogate Chinook salmon bycatch rate (fish per metric ton of pollock) for an unobserved vessel and expands the bycatch rate by the trip delivery weight (metric tons) of pollock to estimate the trip Chinook salmon bycatch. The surrogate bycatch rate for an unobserved vessel trip is apparently computed as a ratio of Chinook salmon bycatch for certain observed hauls to associated weight of pollock caught. Although we cannot find clear documentation for this, apparently ad hoc rules determine which observed vessels and their haul bycatch rate information to use for extrapolation to unobserved vessels. The rules match unobserved and observed vessels by a number of criteria including processing sector, week ending date, gear, and federal reporting area. Apparently, estimation performance by the Catch

Accounting System has not been evaluated, nor is the precision of its Chinook salmon bycatch estimates available.

SAMPLE SELECTION OF INDIVIDUAL SPECIMENS FOR STOCK AND AGE COMPOSITION ESTIMATES

As previously mentioned, the algorithms that will be used to estimate stock composition of the Chinook salmon bycatch (see "Genetics-Based Information on Stock Origins of the Bycatch" and the related Appendix C) rely on the assumption that the collections of scales and tissues for genotyping by the genetics laboratory come from simple random sampling of a theoretical statistical population mixture of individual fish from the contributing stocks. Basic statistical techniques, such as the use of confidence intervals, are also based on an assumption of a random sample from this mixture. A key point is that the sampling rate among hauls, deliveries, or vessels must be approximately the same if the sample of tissues sent to the genetics laboratory is to be considered even approximately a random sample. That means that all sizes and age classes that were present in the bycatch must be included in the sample that is sent to the genetics laboratory at approximately the rate that they occur in the statistical population of salmon in the bycatch. If so, then the probability of inclusion in the sample sent to the genetics laboratory must be approximately uniform and the same for every haul, every delivery, and every vessel.

Individuals of the same source population and age may occur in closer geographic proximity in the ocean than individuals of different sources (McKinnell et al. 1997, Habicht et al. 2006). If this clumping of biological populations occurs at fine geographic scale, the population composition of salmon entering the trawl could trend even during a single haul. If tissue samples were then collected from clusters of salmon within the haul—rather than from salmon dispersed throughout the haul—an insidious and predictable effect on the apparent precision of the estimates would result. That is, because these cluster samples would contain fish more alike in their sources than in the haul as a whole, the estimate of stock composition could be very uncertain and in large error while its estimated precision would appear high. The simple random mixture samples are an important assumption underpinning the genetic mixture algorithms, and this assumption especially affects measures of precision, such as confidence intervals for the stock composition estimates.

Chinook salmon could mechanically sort by size through the process of fishing, trawl retrieval, and catch processing. Within a given mixture of Chinook salmon, age is probably strongly related to stock of origin. Therefore, sampling instructions that do not ensure that individual Chinook salmon are included in the sample with approximately the same probability throughout the bycatch (for each vessel delivery or haul) may produce strongly flawed results. For example sampling instructions that give observers a quota (such as 50 specimens for each delivery), without instructions that result in dispersing the collection throughout the haul, could lead to sample collections that would be far from representative. Again, because size and age almost surely are highly related to stock of origin, any sampling procedure that could favor size or age should definitely not be allowed.

Moreover, if the observer's workload affected the *rate* at which specimens were collected—so that the sampling rate was not standardized across hauls, deliveries, and vessels—then the pooled collection from multiple boats or hauls would not be representative of the catch mixture from the various boats or hauls combined. For example, suppose there were only two boats in the fishery. Further, suppose the bycatch of the first boat was 200 Chinook salmon all of Asian origin, and

the bycatch of the second boat was 50 Chinook salmon all of North American origin. If an observer took 20 specimens from the first boat because of workload or other reasons, but 40 specimens from the second boat, then the ratio in the combined sample would be 33% Asian salmon—quite different from the correct ratio of 80% Asian salmon. Similarly, if a quota of 50 specimens from each vessel were to be collected, the final sample would have 50% Asian salmon, rather than the correct ratio of 80%. Note that any sample that takes specimens at the same rate on both vessels will preserve the correct stock composition ratio in the sample in this simple example. The important points are that the sampling must, at least approximately, result in a random sample at the level of the entire domain or statistical population, even if the statistical population is dispersed within hauls, deliveries, or vessels; the sampling rate must be approximately standardized across hauls, deliveries, and vessels; and most importantly, there must be some definite mechanism to prohibit size selective sampling.

Sampling instructions based on a quota (such as 50 specimens from each delivery) will be wasteful and involve unnecessary expense for collection, processing, and handling, even if the collections for each delivery or haul can be made representative of that haul or delivery. That is, if the overall sample for an entire sampling domain is to be representative, the sampling rate will need to be approximately the same across hauls, deliveries, and vessels. Once the individual specimens have been collected, the only way to ensure that the sample is proportionally representative is to cull and discard specimens from deliveries or hauls with higher sampling rates. To be concrete, suppose the samplers were given instructions to collect 50 Chinook salmon per delivery. For simplicity, assume that the domain included only some number of deliveries or hauls with individual bycatches of exactly 50 Chinook salmon and some other number of deliveries or hauls with individual bycatches of exactly 500 Chinook salmon. Then the usable sampling rate is the lower one: 50/500 or 10%. However, the Observer Program would have had to pay to collect, store, and account for specimens at a 100% sampling rate on those vessels with the lower bycatch. Additionally, the genetics laboratory will have to go through some random culling process, at additional expense, to select only 10% of the samples from the vessels with the lower bycatch numbers (and higher sampling rate). If the genetics laboratory fails to take this step, for the above example, vessels with a lower bycatch will be over represented in the sample. Moreover, there could be something in the style of fishing or gear retrieval that would influence not only their bycatch numbers but also the stock mixture on these vessels. There are many ways that stock mixtures could be related to a vessel's bycatch, especially if fishing techniques were selective for size or age.

A *systematic sample* is one that is found by placing a statistical population in some order, taking a random starting point, and then including every *k*th sampling unit in the sample (Cochran 1977; Thompson 2002). This type of sampling approach has the following advantages: (1) the sample is operationally easier to obtain without mistakes or confusion than a random sample; (2) parameter estimates based on this kind of sampling approach are sometimes more precise than those obtained with random sampling; and (3) there is an obvious assurance that the sample will, in some sense, be representative of the statistical population as a whole. Cochran (1977, section 8.6) notes that in cases of a trend, "systematic sampling is more effective than random sampling", and he notes that if the population is in random order, a systematic sample is equivalent to a random sample (Cochran 1977, section 8.5).

If tissue specimens are to be collected systematically from the Chinook salmon at a high enough sampling rate from fish passing through the observed processing channels, and if that sampling

rate is held constant among deliveries or hauls, then two additional important objectives would easily be met. First, vessel trips, deliveries, or hauls could be grouped into sampling domains based on time and area characteristics without complex calculations of vessel- or haul-specific sampling rates and without a culling process at the genetics laboratory. All individual specimens from the entire fishery would have essentially an equal probability of being included in the larger genetic sample. Second, by randomly sub-sampling from the systematic sample, the resulting set of specimens would very closely approximate a true random sample from the statistical population of all Chinook salmon in the bycatch in that reporting domain, if that were to be important for an analysis.

If the sampling rate is fixed for all salmon at the time of delivery, then the task of determining the desired sampling rate becomes much easier. There are two considerations. First, the sampling rate needs to be high enough so that if bycatch is very low, a statistically reliable estimate of at least the overall stock makeup is possible. Second, if bycatch is large, the sampling rate must be high enough so that the genetics laboratory will be able to create enough sampling domains to supply managers with detail about the change in stock makeup over time and across areas. Going back to Witherell et al. (2002), we note that the changes to management from 2002 were designed to decrease the bycatch to a level of 29,000[2]. In other words, perhaps a genetic analysis would be unnecessary if the bycatch level dropped much below this level. To be concrete, assume that as long as the bycatch remains above 25,000, managers will require at least one sampling domain. From the discussion of sample size in "Genetics-Based Information on Stock Origins of the Bycatch", we might determine that we need a minimum number of 500 tissue specimens for the genetics laboratory. That means if every Chinook salmon in the bycatch is presented to an observer, to reach the total of 500 tissue specimens required by the genetics laboratory, the observers must collect a tissue sample from at least 2% of the Chinook salmon (500 out of every 25,000) in the bycatch. In other words, the observers need to collect a tissue specimen from 1 out of every 50 fish they encounter. If the study is planned this way and the bycatch level reaches 60,000 (and if the observers reach their objective), then the actual sample size will be 2% of the bycatch, or 1,200 tissue specimens. This sample size can be broken into two groups containing at least 500 specimens, allowing two domains. If the bycatch level reaches 80,000, and if the observers reach their objective, there will be a sample size large enough for three domains. However, if the managers actually require six domains for the bycatch of 80,000, then the genetics laboratory will need a minimum of 6 times 500, or 3,000 tissues collections. To get that level of sampling, the sampling rate would have to be increased to 3,000/80,000. That will require the collection of 1 out of every 27 Chinook salmon in the bycatch.

If the marginal cost of laboratory processing of genetic tissues or scales is high as compared to taking, labeling, and storing specimens, then samplers should collect more specimens than will likely be analyzed, because the excess allows several options. First, the genetic collections can be pooled into flexible domains of interest. Second, circumstances may occur in which some population composition estimates have lower precision than deemed satisfactory, and the shortcoming can be remedied by choosing additional specimens from the excess in appropriate domains. Third, special studies that were not anticipated before the season may be completed with the excess. Our recommendation is for the observers to collect the largest sample size

[2] This number is picked here purely to illustrate the principle. In actual practice, such a number would be a policy issue, not a strictly technical one.

practical, by standardizing the sampling interval among observers preseason, with the interval between specimens as small as resources permit.

To implement this system in practice, each observer should be given a random number between 1 and the interval between collections (e.g. between 1 and 50 if the observer is to take a specimen at every 50 Chinook salmon). Then that observer should count salmon up to the random starting point, and then be instructed to continue to collect scales and tissue from specimens at that interval, without regard to vessel, delivery, haul, or processing station. Even if the observer has incomplete information on time and location of catch, a genetic tissue sample should be collected. For example, an observer might be given a sampling interval of 50 and the starting number of 27. The observer would be instructed to take a tissue sample from the 27th Chinook salmon in the bycatch. Then that observer would start counting over again with Fish 1 as the next fish. For an interval of 50, the observer would take a new tissue specimen every 50th salmon after the first (Fish 50, Fish 100, Fish 150, etc.). If that observer's running total was 153 for the year at some point in time, and the observer began sampling a delivery with 42 Chinook salmon on board, the sampler would not collect any tissue specimens from this delivery because the observer would end on Fish 195. If the next delivery had 55 Chinook salmon in it, the observer would collect two specimens from that delivery (Fish 200 and Fish 250).

The actual tissue specimens will consist of the axillary process and five scales from each fish in the sample. These tissues will be placed in an envelope and frozen until processed. Information including the vessel name or plant name, date, fish sequence number, and fish's maturity will be recorded on the envelope. The DNA data will be obtained by staff of the Ted Stevens Marine Research Institute's genetics laboratory from the axillary process of each selected fish, and compared to DNA data from scales, if possible[3]. These scale specimens will be used to determine age of sampled individuals, and then estimate the age distribution of the bycatch.

GENETICS-BASED INFORMATION ON STOCK ORIGINS OF THE BYCATCH

Using artificial marks, such as coded wire tags, the source population of a marked fish can be determined without error. Such information can be invaluable to understanding complex mixed-stock fisheries when large numbers of individuals are marked among at least some of the source populations. For example, Chinook salmon released from many U.S. and Canadian hatcheries are coded wire tagged each year. Recoveries of these tagged fish in coastal fisheries from Southeast Alaska to California are key data used by the Pacific Salmon Commission to infer abundance, survival, exploitation patterns, contributions to the fisheries, and effectiveness of management actions for each of many component populations (Hankin et al. 2005). Similarly, annual patterns on scales and hard parts are commonly used as natural marks by which the sources of fish in catches are identified, although with less certainty than from artificial marks. The genotype of a fish is also a natural mark by which to attempt identification of its source population. As for scale patterns, these genetic marks differ from artificial marks because the identification of a fish's source population can only be done in a probabilistic sense with misidentification usually a risk. Despite their lack of positive source identification, natural marks have important advantages over artificial marks in that all individuals of all source populations carry the marks. Application

[3] Personal communication, Martin Loefflad and Lisa Thompson, NMFS/AFSC.

and recovery of artificial marks is usually so costly that not every individual of a population can be marked, and generally not all populations potentially in a catch mixture are marked.

Genetic marks have become the standard for assessing the population composition of salmon mixtures, especially when numerous source populations may be present. Genetic characters have the important advantage of long-term stability across generations of salmon as compared to natural annual patterns seen on scales and other hard parts that can also be used for population mixture analysis. As a result, the source populations only need to be sampled for genetic characters initially and then infrequently in future years to check for change, whereas the patterns on scales and hard parts generally require annual sampling due to their greater short-term variation.

The genotypes of fish are discrete characters and are passed with DNA molecules from parents to offspring by known rules of inheritance. The DNA occurs in organelles of the fish's cells, either in the nuclei or the mitochondria. The nucleus of each cell of an individual contains the DNA contributed by both parents. The nuclear DNA of a fish is divided into linear segments called *chromosomes*, and the location of a genetic marker on a chromosome is called its *locus*. Each nuclear genetic marker of a fish is determined by a single pair of *alleles*—one from each parent—at the locus. The locus itself has an associated list of the different alleles, or markers, that occur among the individuals comprising the possible bycatch source populations. The number of distinct alleles in the lists varies among the loci. The total number of possible multilocus nuclear DNA genotypes equals the product of the numbers of possible genotypes at the individual loci. Even with only 2 alleles per locus, the number of possible distinct multilocus genotypes for 10 nuclear DNA loci equals 3^{10}= 59,049. As the number of loci and alleles increases, the number of possible multilocus genotypes quickly becomes extremely large.

The mitochondrial DNA is passed as a unit called a haplotype from the female parent to her offspring. Recombination by which offspring can have different combinations of alleles than either of their parents is not operative as it is for nuclear DNA. Therefore, all her offspring will share the female's exact haplotype if no mutations occur. No matter how it is characterized, mitochondrial DNA represents only a single locus, which could limit its utility unless used together with nuclear DNA to identify numerous stocks in their mixtures. The maximum number of populations that can potentially be distinguished in the bycatch from mitochondrial DNA equals the number of distinct haplotypes present among the populations. To use mitochondrial DNA for stock mixture analysis, the relative frequencies of the haplotypes must be estimated or inferred from counts of the different haplotypes in samples from the populations. If the number of distinct haplotypes is large, many may be rare or in low abundance in populations and so would require very large samples to detect their presence and estimate their relative frequencies. Nonetheless, if regional geographic contributions to bycatch are the focus of study, mitochondrial DNA alone can be quite effective (e.g., see Sato et al. 2007).

Both kinds of DNA are used in genetic studies of fish (see introductions by Magoulas 2005, Koljonen and Wilmot 2005, and Wirgin and Waldman 2005). In particular, numerous studies of Pacific salmon populations using nuclear and mitochondrial DNA have been performed, and the potential for these genetic marks in identifying the sources of the salmon bycatch continues to be evaluated. The first applications to identification of salmon origins in the North Pacific used allozymes, which are variant forms of enzymes coded by different alleles at nuclear DNA loci. Allozyme genotypes are made observable in the laboratory through electrophoresis in which the charged protein molecules are separated in an electric field. Allozymes were used to estimate the

stock composition of Chinook, sockeye and chum salmon in the illegal catch of the vessel *Arctic Wind* seized south of the Aleutian Islands in 2000 (Wilmot et al. 2000). The Chinook salmon stock composition was estimated using an allozymes baseline of 253 populations and 22 loci from the Pacific Rim (Teel et al. 1999) together with three additional Yukon River populations. The chum salmon stock composition was estimated from an allozyme baseline of 273 populations and 20 loci from the Pacific Rim. Among other studies using allozymes to identify sources of salmon in the North Pacific Ocean was an earlier attempt to determine composition of chum salmon bycatch in the pollock fishery with a Pacific Rim allozyme baseline of 77 populations and 20 loci (Wilmot et al. 1998). The amount of variation available from allozymes is somewhat limiting, multiple tissues are often required, and, importantly, enzymes begin to degrade after death of an organism so tissues must be frozen as soon as possible. Therefore, allozymes have fallen out of favor to nuclear and mitochondrial DNA genetic marks that provide far greater variation and require less tissue and care in its storage. The two main nuclear DNA marks in use today are called microsatellites and SNPs.

Both nuclear SNPs and microsatellites can only be observed using biochemical techniques in the laboratory, but their assays for large numbers of individual fish are possible with throughput rapid enough even for inseason analyses. SNPs have an advantage over microsatellites in that interpretation of genotypes among different laboratories is much simpler and the potential for laboratory automation appears greater. Almost all common SNPs have just two alleles, whereas microsatellites can have various numbers of alleles ranging from a few to 50 or more. The identification of SNP and microsatellite alleles for an individual salmon has become automated, resulting in lower cost, faster assays, and lower error rates than when performed manually by laboratory personnel (Seeb, et al. 2009). Advances in automation continue to the present with Garvin and Gharrett (2009) recently inventing a powerful method for surveying populations to discover useful SNPs rapidly and inexpensively, thereby adding further to the appeal of SNPs. Further, the authors improved a genotyping assay to screen SNPs in thousands of fish cheaply relative to standard commercial assay, and also improved on resolution of the phase of nearby SNPs in the same gene.

The list of a fish's single-locus genotypes, i.e. its allele pair at each assayed locus, is called the fish's *multilocus genotype*. Ordinarily, each SNP has just three types possible for an individual fish at a particular locus, whereas some microsatellites can comprise hundreds or even thousands of different genotypes at a locus. The number of multilocus genotypes possible can be very large because it equals the product of the numbers of single-locus genotypes possible at the comprising loci. Of course, a single SNP is not as informative regarding population sources as is a single microsatellite, so more individual SNP loci are required than microsatellite loci to achieve the same level of precision in population mixture analysis. However, recent work by Smith and Seeb (2008) demonstrated that the ability to choose SNP alleles for information content mitigated the effect of fewer genotypes being available. The larger number of SNP loci assayed may be somewhat concerning because the chances become greater that some will be physically linked on the same chromosome and might be in what is called *linkage disequilibrium*. If any of the SNPs are in linkage disequilibrium, the computation of multilocus relative frequencies in the source populations required for population mixture analysis could be more difficult, as explained in Appendix C. Testing for disequilibrium between locus pairs is a standard practice in surveys of genetic markers and can easily be accounted for prior to use in mixed stock analysis by either dropping one of the linked loci or by combining the genotypes into a single locus which is treated similar to mtDNA.

Baseline samples for both SNPs and microsatellites from many of the possible source populations of Chinook and chum salmon have already been obtained in other studies. If those samples prove adequate, the main cost for source population analysis of the bycatch is from collecting and laboratory processing of the bycatch tissue specimens. The Genetic Analysis of Pacific Salmonids consortium of nine major West Coast salmon genetics laboratories created the main microsatellite baseline dataset for Chinook salmon (Moran et al. 2005, 2006, Seeb et al. 2007). Currently this database contains information for 13 loci from 22,269 fish in 304 samples of 166 populations among 44 reporting regions ranging from southeast Alaska to the Central Valley of California (Moran et al. 2006). Before this baseline can be used for the Chinook salmon bycatch analysis, it will either need to be extended to west and north of Yakutat, Alaska, to include all of Alaska and Russia, or else used in conjunction with another baseline of broader geographic coverage.

In 2008, the SNP baseline for Chinook salmon developed by the Alaska Department of Fish and Game included 53 loci from 23,300 individuals of 172 populations ranging from California to Russia (Templin et al. *in press*). This baseline has been used for estimating stock proportions in the bycatch, in juvenile samples from the Bering Sea, and in several fisheries throughout Alaska[4]

A Pacific Rim chum salmon baseline of 13 microsatellite loci has been developed from over 34,000 chum salmon sampled at 283 locations from Japan, Russia, Alaska, British Columbia, and Washington (Beacham et al. 2005). A study of simulated mixtures showed the potential of the baseline to provide reliable estimates of regional stock composition even when the entire Pacific Rim populations might contribute. Gharrett et al. (2007) developed another chum salmon baseline of 12 microsatellite loci for 4,900 chum salmon that were sampled from 73 populations distributed from Puget Sound to Honshu Island and Primoriye in Asia. Because some overlapping information with other baselines already existed for many western Alaskan populations, the focus was on Asian populations as well as populations from Southeast Alaska, British Columbia, and Puget Sound. The Alaska Department of Fish and Game has developed a SNP baseline for chum salmon which includes 61 loci and 165 populations ranging from the Pacific Northwest to Russia and Japan (Seeb et al. *in press,* Alaska Department of Fish and Game *unpublished data*). Simulation studies show that these baselines can provide a basis for estimating regional compositions of Bering Sea fisheries chum salmon bycatches. Improving the baseline could be achieved by increasing sample sizes, surveying additional populations, and judiciously increasing the number of loci used, particularly ones that will increase resolution of coastal western Alaska stocks.

Although the ideal solution to the geographic questions would identify the bycatch sources to the individual contributing populations; that level of detail is not yet possible with current levels of sampling or genetic diversity. Commonly, populations from nearby streams are more similar genetically, i.e., have smaller differences in relative frequencies of the possible genotypes, than those of distant streams. Of course, the more similar stocks are, the more difficult to separate their individuals in catch mixtures to sources or to estimate the population proportions composing the catch. Instead, the coast-wide populations of Chinook and chum salmon are combined into regional reporting groups in order to provide more accurate and precise estimation of larger groups. The possibility to employ broad scale geographic baselines to separate bycatch

[4] Personal communication, Bill Templin, Alaska Department of Fish and Game.

to regional origins, followed by application of a local baseline to the same tissue samples to refine separation for a particular region does not seem to have been explored. Users have local interests and inevitably request finer stock resolution. The potential savings from less baseline development should encourage this focus for separation within regions as more genetic loci are discovered.

UNCERTAINTY AND SAMPLE SIZE

The uncertainty in the estimates of stock composition of the bycatch results from sampling, or examining only a portion of the bycatch and from the imperfect identification of population sources of the individuals in that portion using the genetic characters (Kalinowski 2004). The uncertainty from simple random sampling of the bycatch can be approximated using the multinomial probability model. The bycatch has an unknown fixed source population composition, but the potential bycatch samples will vary in their source compositions. Therefore, even if use of the genetic characters correctly identified the source of each sampled fish, any estimate of the source population composition based on the sample would vary about the actual source composition of the bycatch. However, this variation becomes smaller as the sample size is increased, providing a means to control it. The minimum bycatch sample size for a population mixture analysis required to control bycatch sampling variation can be determined from the simultaneous confidence intervals theorem of Fitzpatrick and Scott (1987) and Thompson (1987) for multinomial proportions (see also pp. 42–44 of Thompson 2002). Two variables describe the degree of control, and they must be specified in order to determine the minimum required sample size. The first control variable, denoted by α, is the maximum probability that a bycatch sample has one of the source population proportions unacceptably different than the true source population proportion. The second control variable, denoted by d, is the maximum difference allowed between any of the estimated and true source population proportions. For example, a sample of $n = 403$ fish is required to be 90% ($= (1-\alpha)\cdot 100\%$) confident that none of the source population proportions in the bycatch sample differ by more than 0.05 from the corresponding proportions in the bycatch itself (refer to Table 5.1, p. 43, Thompson 2002, after setting $\alpha = 0.10$ and $d = 0.05$). The assumption underlying this computation is random sampling of the bycatch, and if systematic sampling is used instead, the required sample size may be smaller (because variation of systematic sampling is likely smaller than random sampling), but how much less is unknown. Although the sampling variation of the source composition has been controlled by the minimum sample size found from the theorem, the final source composition estimates may be more variable than desired, as discussed next.

The other cause for uncertainty in the source population composition of the bycatch is the imperfect identification of the source populations of bycatch individuals using the genetic characters. If the conditional maximum likelihood method for estimation of the source population composition of the bycatch (Fournier et al. 1984; Millar 1987; Pella and Milner 1987) is used, this uncertainty can only be reduced by either increasing the number of informative loci that define the multilocus genotypes, or by increasing the baseline sample sizes to better estimate the relative frequencies of alleles at loci among the source populations, or both. On the other hand, if the Bayesian method of source population composition estimation (Pella and Masuda 2001) is used, this uncertainty is reduced to some extent by making better use of the bycatch sample than does likelihood estimation. Specifically, the Bayesian method extracts information in the bycatch sample about the relative frequencies of alleles at loci for the source populations present.

The best identification or assignment rule is based on the so-called posterior source probabilities of an individual (Pella and Milner 1987, Pella and Masuda 2001). For now, the posterior probability that an individual originated from a particular source population is a function of the source-population composition of the entire bycatch and the relative frequencies of the individual's multilocus genotype in each possible source population. The relative frequency of the genotype in a source population is a function of that population's unknown relative frequencies of its alleles at each locus, and estimates from the baseline samples (possibly enhanced with information from the mixture sample) must be substituted for these unknowns. Increasing the sample sizes of alleles at loci from the source populations will reduce errors in estimated genotype relative frequencies, and concomitantly in source assignments. However, even with the actual genotype relative frequencies, some source assignments will be in error because the various posterior source probabilities never equal 0 or 1 as would be necessary for certainty and error-free assignments. The posterior source probabilities will approach these ideal values, 0 or 1, as more informative loci are added and baseline samples are increased. With few genetic loci, the posterior source probabilities of many mixture individuals may be nearly equal for several of the baseline populations.

A Simple Example Mixture Analysis Using Colored Balls and Urns

Next, the principles underlying the identification of sources of individuals in the bycatch will be illustrated with a simple example involving colored balls and urns. We will demonstrate how differences in the relative frequencies of universally shared types can be used to estimate the source composition of a mixture from a sample. The colored balls (red, white, and blue) of 3 urns are combined into a single mixture urn (Table 4). The first urn contained 1,600 red, 200 white, and 200 blue balls for a total of 2,000 balls; the second urn contained 350 red, 6,300 white, and 350 blue balls for a total of 7,000 balls, and the third urn contained 100 red, 200 white, and 700 blue balls, for a total of 1,000 balls. The combined mixture urn contains a total of 10,000 balls with color composition of 2,050 red, 6,700 white, and 1,250 blue balls.

Table 4.–Numbers of colored balls in 3 source urns and the combined mixture urn.

Color	Urn 1	Urn 2	Urn 3	Mixture urn
Red	1,600	350	100	2,050
White	200	6,300	200	6,700
Blue	200	350	700	1,250
Total	2,000	7,000	1,000	10,000

If the mixture urn is well shaken and a single ball is drawn from it, what is the probability that it is from the first urn? The answer is 0.2 (= 2,000/10,000), the proportion of the mixture from urn 1. This probability is called the *prior source probability* and will be denoted by $p(1)$. The term *prior* is used because we are not told its color yet, i.e., the *data* from the observation is not available. The prior source probabilities for the second and third urns are $p(2) = 0.7$ (= 7,000/10,000) and $p(3) = 0.1$ (= 1,000/10,000), respectively. Notice that the sum of the prior probabilities is 1 because the three urns are the only contributors. What is the relative frequency of red balls in the first urn? The answer is 0.8 (= 1,600/2,000), the proportion of balls in the first urn that are red, which will be denoted as $f(red \mid 1)$. (Later, analogous functions will be defined for the relative frequencies of genotypes in source populations.) The relative frequencies of red

balls in the second and third urns are f (*red* | 2) = 0.05 (=350/7,000) and f (*red* | 3) = 0.1 (= 100/1,000), respectively. What is the probability that the color of the ball drawn from the mixture urn is red? The answer is 0.205 (= 2,050/10,000), the proportion of the mixture urn balls that are red. This probability will be denoted as θ_{red}, which can also be computed as $\theta_{red} = \sum_{i=1}^{3} p(i) f(red \mid i)$. The probabilities that the color is white or blue are $\theta_{white} = \sum_{i=1}^{3} p(i) f(white \mid i)$=0.670 (= 6,700/10,000) and $\theta_{blue} = \sum_{i=1}^{3} p(i) f(blue \mid i)$=0.125 (= 1,250/10,000), respectively. If the color of the ball drawn from the mixture urn is red, what is the probability that its source was the first urn? The answer is 0.780 (= 1,600/2,050), the proportion of red balls in the mixture urn contributed by the first urn. This probability is called the *posterior source probability* of the first urn given the ball was red, it can be denoted as P(1| red), and can also be computed as $P(1 \mid red) = p(1) f(red \mid 1) \Big/ \sum_{i=1}^{3} p(i) f(red \mid i)$. The term *posterior* is used because we are now told the color of the ball, i.e., the *data* have become available, whereas the term *prior* was used before we knew the color. The posterior source probabilities for the second and third urns given a red ball are P(2| red) = 0.171 (= 350/2,050) and P(3| red) = 0.049 (= 100/2,050). Notice that the sum of the posterior source probabilities for a red ball is 1. If the source of the red ball were needed, the best guess (lowest error rate) would be the first urn for which the posterior source probability (0.78) is greatest. This rule is called the *maximum a posteriori* (MAP) rule. (Some geneticists assign individuals by their genotypes to populations based only on the relative frequencies of the genotypes in the possible source populations. This rule is very successful if the mixture composition is uniform among all the sources, but it misleading and optimistic if the uniform mixture composition is assumed to be unknown.)

In direct analogy to genetic population mixture analysis is the problem of estimating the proportions contributed by each of the urns to their mixture, i.e., the prior source probabilities, $\mathbf{p}=(p_1, p_2, p_3)'$, when the numbers of balls per urn are unknown. We assume that the relative frequencies of the colors in each urn are known and that a sample of balls is drawn from the mixture urn and their colors are recorded. (Later the estimation methods will allow the relative frequencies of the colors in each urn to be unknown if estimable from a sample.) A sample of 100 balls is drawn with replacement (this simplifies computations due to the finite numbers of balls per urn) from the mixture urn and the numbers of red (n_{red}), white (n_{white}), and blue (n_{blue}) balls are recorded. (Notice that the expected numbers of red, white, and blue balls equals $100 \cdot \boldsymbol{\theta}$ = 20.5, 67.0, and 12.5, respectively, where $\theta = (\theta_{red}, \theta_{white}, \theta_{blue})$). The probability function for the counts is the familiar multinomial function, which provides the probabilities of the counts,

$$\Pr(n_{red}, n_{white}, n_{blue}) \propto \theta_{red}^{n_{red}} \theta_{white}^{n_{white}} \theta_{blue}^{n_{blue}} .$$

The unknown mixture proportions from the urns can be made explicit in this expression by substituting definitions of the θ's, namely,

$$\Pr(n_{red}, n_{white}, n_{blue}) \propto \left(\sum_{i=1}^{3} p(i) f(red \mid i) \right)^{n_{red}} \left(\sum_{i=1}^{3} p(i) f(white \mid i) \right)^{n_{white}} \left(\sum_{i=1}^{3} p(i) f(blue \mid i) \right)^{n_{blue}} .$$

The likelihood function for **p** equals the probability function, but now viewed as a function **p** and the mixture sample counts viewed as fixed known constants. We show below that the likelihood function can be easily maximized to provide an estimate of **p**. An alternative and seemingly superior approach is to use the multinomial probability function in a Bayesian estimation method for **p**. In real world population mixture analyses, the numbers of populations (here urns) can be in the hundreds and the numbers of genotypes (colors) can be an extremely large number.

Although the urn example is simple, the statistical estimation of population composition of a 3-population mixture by scale pattern analysis follows it closely. An assignment rule is developed for the scale patterns, which divides the individuals of each population (like an urn) into 3 classes (like the colors). The rule is applied to a sample of the mixture and to a large sample from each of the source populations. The counts of individuals assigned to each population from each sample are used to estimate the mixture composition. Typically, the rule is designed to assign most individuals of a population to one class and as few as possible to the other two classes. Originally called the classification with correction method, it can now be seen to be better resolved as a special case of the maximum likelihood and Bayesian estimation methods discussed in detail in the Appendix C (also see Millar 1987; Wood et al. 1987; Pella and Masuda 2005). Also in the Appendix C is a detailed explanation of the general problem of estimating the source population composition of the bycatch from genotypes observed in a mixture sample.

RECOMMENDATIONS

Before making these recommendations, we reviewed written plans and manuals from the National Marine Fisheries Service dealing with the Observer Program, and we spoke with some of the people involved with the operation of that program. We reviewed published estimates derived from data collected by this program, and tried to review written explanations of how the estimates were generated. We also reviewed Miller's (2005) proposed statistical estimation methods, and tried to compare these with what the National Marine Fishery Service is currently using. We reviewed the instructions observers are given on how to make specific tissue collections, and we also reviewed basic statistical references to see if we could find guidance for improving these collections. Finally, we reviewed the multiple objectives of the Observer Program, as these relate to salmon bycatch in the Bering Sea pollock fishery. Broadly speaking, the problem of generating estimates of salmon bycatch by its regional geographic origins requires two components. The first component involves estimating the size of the bycatch for seasons, shorter periods, and specific areas, collectively called domains. The second component involves generating a genetics-based analysis of the stock mixtures so that the bycatch estimates can be expressed in terms of its regional geographic origins. Our recommendations, therefore, will be divided somewhat artificially along the line of these two components.

ESTIMATION OF DOMAIN BYCATCH

At the end of our review we still do not completely understand the methods that the National Marine Fisheries Service used to estimate salmon bycatch, and there are discrepancies between their estimates and other sources. An unanswered question remains: could the discrepancies be partially explained by sampling error? The National Marine Fisheries Service did not publish measures of sampling error, such as confidence intervals, but the current regulations seem to allow for measuring salmon bycatch using a total count of salmon in the bycatch, so that sampling error should not be an issue at all. Under the regulations, all of the Chinook salmon in the bycatch of the Bering Sea–Aleutian Islands pollock fishery must be delivered to the

motherships, catcher-processors, and shore-based and floating processing plants. Observers are stationed at all these delivery locations so a bycatch census should be possible. If so, we certainly recommend producing and publishing these census counts, even if only for the purpose of comparison.

Two observers are present on each mothership and each catcher-processor, and combined they can, in principle, effectively count each Chinook salmon of every haul delivered to the motherships or caught by the catcher-processors, although in practice other duties may conflict with achieving this goal. The catcher vessels delivering to the motherships deliver the unsorted catch in the trawl codend, so no additional observer is needed on these catcher vessels. One observer is always present on the large catcher vessels that deliver to shore-based and floating plants, and that observer can, in principle, also count each Chinook salmon of every haul if the catch is sorted at sea. If the catch is not sorted at sea, the Chinook salmon for all hauls are required to be stored and delivered to the processing plant where an observer can count each fish. Also, plant personnel record the count of Chinook salmon in the deliveries. Therefore, the bycatch by every large catcher vessel is always under an observer's eyes and each Chinook salmon of either each haul, or else the combined hauls of each vessel delivery, can be counted.

The midsized catcher vessels that make their deliveries to shore-based or floating processing plants are the only catcher vessels for which an observer is not required on board during 100% of the fishing days. An observer is required to be on board the midsized catcher vessels for 30% of fishing days, so during these days with an observer onboard, each Chinook salmon can be counted as was the case for the large catcher vessels. For the 70% of fishing days when an observer is not on board midsized catcher vessels, the vessel is required to store the Chinook salmon for delivery to the plant. During the 70% of days without an onboard observer, the bycatch is not under the eyes of an observer. The potential for uncounted discard of Chinook salmon bycatch at sea exists if the regulations are violated. To assure that discard does not occur, several checks should be used.

First, a statistical test for the midsized vessels could be performed that compares their vessel-specific pairs of average bycatch rates (number of Chinook salmon caught per ton of pollock) during each quarter for trips with and without an observer onboard. If the presence of an observer does not affect the bycatch rate, the probability that the number of vessels for which the average bycatch rate for trips with an observer is greater than the average bycatch rate for trips without an observer can be evaluated with the binomial probability function with occurrence probability equal to one-half. Above, we described this as a two-sided test for the sake of completeness. It seems more realistic that there would be little or no benefit to the vessel for a change in fishing practice or change in fishing location when an observer was on board. One could doubt that the vessels would devote 30% of their fishing days to a practice that lost opportunity for pollock catch, whereas discarding salmon during the 70% of the trips without observers may seem like a more credible possibility. If these tests were formulated as one-sided tests, they would be more powerful as tests of the observer effect than the two-sided test. If nothing is amiss, on average, half of vessels will by chance have a higher average rate when an observer is onboard, and half of vessels with an observer onboard will have a lower rate. Lower and upper bounds can be found for the numbers of vessels with a higher bycatch rate when an observer was onboard such that the probability of more extreme counts under pure chance is known exactly. Comparing the counts of vessels to these bounds provides any evidence of salmon dumping when no observer was present (the observer effect). For instance, the bounds

might be set so that the probability that pure chance could account for the count falling beyond the bounds was less than 0.1%. Then, if the count did fall beyond the bounds, an investigation into the source of the discrepancy would be needed.

Assuming that the test result is consistent with the vessels following regulations (i.e., pure chance can account for the differences in bycatch rates for trips with and without an observer) the bycatch counts at the plants for midsized vessels without an observer can be viewed as the actual bycatch values. Then the counts of Chinook salmon in the bycatch can be considered to be a census and there would be little legitimate reason to question the official estimates. However, the test may not be so clear when the count fell into a range that would be considered "somewhat unusual," e.g., 0.1 or 0.2. Because this test could be repeated quarterly, following these results over time should completely settle the question. However, these tests can only be performed if the National Marine Fisheries Service commits to a complete and accurate accounting of all Chinook salmon presented to observers, irrespective of whether or not an observer was on the vessel during the fishing trip.

There will always be missing observations of salmon counts, whether because of workload, loss of data, data quality questions, or other reasons. Miller's (2005) model-based estimation for undersampled trips can be used after careful consideration of the sampling situation, and this approach will also allow the public to know exactly how the estimates were generated. This method will allow a reasonable estimate of sampling uncertainty with the estimate. Miller's methods appear closely related to the method currently used by National Marine Fisheries Service Alaska Regional Office, as far as we can determine.

The pollock catch of any vessel trip is known and can be used in estimating the corresponding Chinook salmon bycatch provided an estimate of the Chinook salmon bycatch rate (number of Chinook salmon per metric ton of pollock) is available. The problem of estimating the bycatch rate for a vessel trip without a determination of the bycatch (or for trips without an observer on board) remains open, and here we suggest possibilities. First, vessel-specific data on bycatch rates from trips with an observer onboard can be used. The estimate of bycatch during trips without an observer onboard the vessel should be set equal to the product of the vessel's bycatch rate from trips with an observer and its pollock catch. This approach does not utilize any information about the domains in which the vessel fished, but is simple and should always be possible after the vessel has made at least one trip with an observer onboard. If domain-specific data is to be used, the vessel bycatch rate for unobserved trips could be estimated from bycatch rates of itself and other vessels fishing in the domain that had observers on board. The critical decisions are (1) which vessels and trips with an observer onboard to include in estimation of the bycatch rate, and (2) how to combine this information. Possibly a kernel estimator of bycatch rate similar to that for the Japanese squid driftnet fishery (Pella et al. 1993) would apply and simplify the decisions: a weighted average of bycatch rates from other vessel trips with an observer could be used with weights depending on their geographic and temporal proximity to the vessel for which the bycatch rate is to be estimated. The values for the weights would be chosen to minimize a function of cross-validation prediction errors.

GENETIC ANALYSIS

Fishery samples

The Chinook salmon bycatch of the entire Bering Sea–Aleutian Islands trawl fishery should be sampled systematically for individual fish, from which genetic tissue specimens and associated

scales for aging are drawn. This approach is operationally simple and easy to understand, this will definitely produce a representative sample, and this approach will be less wasteful and more efficient than approaches that allow the sampling rate to vary widely among vessels, trips, or hauls. To implement this system in practice, each observer should be given a random number between 1 and the interval between collections (e.g. between 1 and 50 if the observer is to take a specimen at every 50 Chinook salmon). Then that observer should count salmon up to the random starting point, and thereafter continue to collect scales and tissue from specimens at that interval, without regard to vessel, delivery, haul, or processing station. Coordination between observers is necessary if more than one observer has an opportunity to sample a catch so as to avoid double sampling. This concern can be easily addressed by establishing simple rules that put one unique observer in charge of tissue collection for each delivery. Even if the observer has incomplete information on time and location of catch, a genetic tissue sample should be collected. Our recommendation is for the observers to collect the largest sample size practical, by making the interval between specimens as small as resources permit, and then by standardizing the sampling interval among observers preseason. As we previously mentioned, by collecting more specimens than will likely be processed in the genetics laboratory, the excess allows the genetic collections to be pooled into flexible domains of interest; circumstances may occur in which some population composition estimates have lower precision than deemed satisfactory, and the shortcoming can be remedied by choosing additional samples from the excess in appropriate domains; and special studies that were not anticipated before the season may be completed with the excess.

Baseline evaluation

The regional stock composition of the bycatch may have important management consequences. If the bycatch is sampled as recommended, the stock composition of the fishery sample will be known to closely approximate that of the entire bycatch. Then, if each individual in the fishery sample could be identified to its source stock without error, the users could have full trust in the apparent bycatch composition based on the sample. However, source identification from genetic information is not yet without error because of the limited number of loci included. Users of the stock composition assessment will want to have a transparent demonstration of how well the genotypes of the contributing stocks can be identified to source. The standard practice for assessing baseline power, or the capacity of a baseline to distinguish among the stocks, is to analyze many sets of simulated baseline and mixture samples using allele relative frequencies observed in the actual baseline samples as though they equaled the corresponding unknown values in the stocks. The hypothetical stock mixture from which the simulated mixture sample is drawn has stock composition specified by the analyst. Each set of simulated samples—baseline samples and a fishery mixture sample—are analyzed using the conditional maximum likelihood method to estimate the stock proportions. Then the distributions of the estimates of stock proportions from the many sets are compared with the known stock proportions of the underlying mixture. This method of baseline evaluation has been used for 20 years or more. Commonly, so-called 100% mixtures are specified in which only a single stock, or stock group, composes the mixture with the remaining stocks absent. The baseline power determined from these 100% simulations likely provides an optimistic view for several reasons. First, the fishery-sampling component of uncertainty is completely missing. If the simulated mixtures included several stocks, the actual stock compositions of simulated mixture samples would vary from that specified. Nonetheless, if the simulated mixture samples are sufficiently large, e.g., sample size is specified using the theorem for simultaneous confidence intervals of multinomial proportions

(Thompson 2002), this effect is negligible. Second, the uncertainty in stock composition increases as the underlying stock proportions are moved away from the constraining boundaries of the simplex, $\mathbf{p} = (p_1,\ldots,p_c)$ where $0 \le p_i \le 1$, $i = 1,\ldots,c$ and $\sum_{i=1}^{c} p_i = 1$. This increase is caused not only by the inclusion of the fishery-sampling component of variation in actual stock composition of samples, but also by confusion among the additional stocks present. The 100% mixtures do provide good information about the bias and variation at the boundary caused by the variable estimates being constrained. However, actual bycatch mixtures in the pollock fishery will usually be composed of several stocks and 100% mixtures will likely understate their uncertainty. Third, and last, recent investigation of the standard simulation method shows that it overstates the predicted accuracy with profound bias for closely related stocks is less biased for reporting groups, but increases as more genetic loci are added to the baseline (Anderson et al. 2008). A cross validation method developed by Anderson et al. (2008) should be used in a reevaluation of baseline power of SNPs and microsatellites and their combination.

In addition to 100% simulations, equal-proportions simulations would provide a less optimistic—possibly a worst case—evaluation. If the mixture sample size is assumed to be large as planned, the simulations can be simplified to generate large numbers of genotypes from each of the stock baseline samples. Then these simulated genotypes are classified to their purported source stocks with the maximum *a posteriori* rule: assign the individual to the baseline stock for which its genotype is most frequent. A so-called confusion matrix can be produced from the assignments, which clearly demonstrates the baseline power to identify stock sources of individuals under this most challenging condition. The confusion matrix shows the stocks that are most difficult to distinguish. The data from this experiment can also be summarized to provide a confusion matrix for regional groups of stocks. Users of stock composition analysis are already familiar with the confusion matrix from discriminant analysis using scales, and therefore the results should be readily appreciated.

A number of genetic baselines have been developed, and progress in improving their geographic coverages and in discovering new and useful genetic markers will continue. Our emphasis on resolving the bycatch origins has been on its regional geographic composition because that appears practical in the near term. Users of the genetic analysis of the bycatch, such as Yukon River subsistence fishers, will probably demand the highest resolution possible. Genetic marker surveys within regions such as the Yukon River have been useful to help satisfy such requests at least when the mixture is composed only of local stocks, and presumably future surveys will provide even better local resolution. Different regions may require their own marker sets to separate local stocks. A complex of genetic baselines can be anticipated, including a Pacific Rim baseline and additional regional baselines. The Pacific Rim baseline would be composed of genetic samples from all the stocks for those loci useful for regional separation. Regional baselines would include samples from the regional stocks with additional genetic loci useful for separating their local stocks. Past statistical treatments of genetic mixture analysis have considered an essentially uniform baseline of shared genetic markers available for all stocks, possibly with a few loci missing for some stocks. With the complex of genetic baselines that may occur for the Bering Sea–Aleutian Islands salmon bycatch, the issue of missing loci in baseline samples as dealt with to this time will likely become more problematic. Statistical methods and computer programs may be required for sequentially resolving the bycatch, first regionally, and second, locally. For example, the regional resolution could be in the form of probabilities of

regional origins of the bycatch individuals. Then the regional markers not yet used could be applied to provide more detail. Although all the markers would have to be assayed for the bycatch tissue samples, a savings would occur by limiting additional baseline development for local detail to useful loci within their regions of high interest.

Finally, the discovery of bias in simulations by Anderson et al. (2008) recalls an early issue of applications of genetics to fish mixtures. The addition of genetic loci to a baseline were purported to either improve, or at worst, not reduce, the success in resolving a stock mixture. That may be true, but if so, it avoids a pitfall sometimes found in discriminant and pattern recognition analysis (Hand, 1981). As additional variables are added to classify generic objects to their sources, the success rate can at first increase, but then decrease. The evaluation of baselines and their loci needs to be alert to this danger.

Bycatch stock mixture analysis

The Bayesian method of estimation is preferred for two reasons. First, the Bayesian method allows the analyst to speak about uncertainty of the bycatch stock composition in terms expected by the non-technical user. Bayesian probability intervals for the unknown stock proportions mean that the unknown stock proportions composing the bycatch lie within the stated intervals with probabilities as given. Frequentist confidence intervals are much more complicated: they refer to the frequency with which the computed interval covers the unknown stock composition **p**. Second, the Bayesian method makes better use of the genetic data by adding the information about allele relative frequencies of the mixture sample to that of the baseline samples. Given the large mixture samples planned, this information could be substantial.

KEY RECOMMENDATIONS

- The internal workings of the Catch Accounting System cannot currently be evaluated, nor can the precision of its Chinook salmon bycatch estimates be calculated. The National Marine Fisheries Service should develop written procedures for the assessment of the quality of the data and resulting estimates, both in terms of sampling errors and non-sampling errors. These written procedures should be written in standard statistical notation and they should be peer reviewed.
- Bycatch numbers could, in principle, be ascertained by a census. This approach is simple, easy to explain, and has the advantage that it is free of sampling error. We recommend that the National Marine Fisheries Service at least consider moving towards a more census-like approach to assessing the bycatch magnitude. Further, a census approach will permit statistical tests to determine that the presence of an observer on board a vessel does not affect the level of bycatch that is reported.
- The National Marine Fisheries Service should develop a strict protocol to ensure representative sampling. In particular, size selective sampling should be absolutely prohibited and sampling rates should be as uniform as possible across hauls, deliveries, and vessels. To achieve this uniformity, we recommend systematic sampling of the Chinook salmon bycatch to obtain tissue samples for mixture analysis. This approach is operationally simple, easy to understand, produces a representative sample, and will be less wasteful and more efficient than approaches that allow the sampling rate to vary widely among vessels, trips, or hauls.

- We offer detailed recommendations for Bayesian genetic mixture analysis and baseline evaluation, and we offer specific recommendations for demonstrating the validity of the method using simulation.

CONCLUSION

Our recommendations for collecting the Chinook salmon tissues samples are quite a departure from the way that Chinook salmon in the bycatch has been sampled in the past. We recognize that it might be impossible to move rapidly towards the ideal of a truly representative sample of Chinook salmon in the entire bycatch. Even so, some departures from this ideal present bigger problems than others.

Any method for collecting genetic tissue specimens that could be size selective in the salmon that are chosen should be definitely rejected. Size selective sampling will almost surely produce a highly biased picture of stock origins and of the level of sampling error (reflected in confidence intervals, standard error estimates, or Bayes probability intervals). Similarly, a very accurate picture of the size and age distribution of salmon in the bycatch will be needed in the future, so that bycatch mortalities can be expressed on a returning adult equivalency. Salmon taken early in life will have many opportunities to fail and die before their final homing migration. The bycatch of a very young fish will have a different and less pronounced effect on the terminal run than salmon caught immediately prior to this homing migration, when the subsequent mortality rate would have been expected to be comparatively low if they had not been caught by the trawl fishery. So, most importantly, the sampling process must involve a definite protocol that assures salmon of all sizes are equally likely to enter the sample for genetic decoding as well for age and size distribution estimates.

Another serious bias in estimation of stock origins of the bycatch would also occur if only catcher vessels carrying an observer and delivering to shore-based and floating processors were sampled for genetic tissues. The availability of an extra observer during their deliveries would seem to more easily enable such sampling than deliveries to motherships and hauls of a catcher-processor when only a single observer is present. The catcher vessels delivering to motherships as well as the catcher-processors may distribute their fishing efforts differently over the Bering Sea–Aleutian Islands area than the catcher vessels delivering to shore-based and floating processors. If so, the stock composition of their salmon bycatch is unlikely to be represented by the sampled bycatch for the catcher vessels delivering to shore-based and floating processors.

If tissue samples are all taken from spatial or temporal clusters of bycatch, this could greatly affect the actual and reported precision (e.g., confidence intervals, standard errors, or Bayes probability intervals) of stock composition. This clustered sampling, for example, would result by taking specimens of only a small subset of vessel deliveries, or from taking specimens from spatially clustered fish within the hauls. If this kind of sampling is the only practical way to collect the specimens, then the problem could be informally addressed in the genetic laboratory by decoding vastly more specimens than Thompson's (1987) sample size algorithm recommends. Even still, with these cluster samples the confidence intervals would be biased. Although this should be a source for concern when considering the estimates for a single domain, in principle, estimates that result from this kind of tissue collection from clusters should follow the correct trends over time and space.

If tissue specimens were collected from a quota, such as 50 specimens from each delivery, this would have an unknown effect on both accuracy and precision. If, for some reason, vessels with higher bycatch rate have a greater proportion of Western Alaskan salmon in the bycatch, then samples based on vessel or delivery quotas will tend to produce estimates of stock composition that will understate the level of Western Alaskan stocks in the total bycatch, because vessels with larger numbers of Western Alaskan salmon will be sampled at a lower rate. If quota-based sample collection is the only practical way to obtain the tissue specimens, then the final sample that is actually decoded should be carefully weighted to achieve as close to a fixed sampling rate as possible.

Like explaining how an automatic timepiece works, explaining how a salmon stock mixture analysis works in principle is a lot easier than actually making one work. In both cases the art of actually doing it is all about making a large number of interrelated pieces work together correctly and precisely. However, in the case of the problem producing stock composition estimates, simply making it work is only part of the larger problem. The entire problem has not been solved until reasonable skeptics can be convinced that it works—that is, that the estimates are credible. We believe our recommendations that deal with proving that the estimates are credible are at least as important as our procedural recommendations for improving the estimates.

ACKNOWLEDGEMENTS

We are grateful to many members of the staffs at the Alaska Department of Fish and Game and the National Marine Fisheries Service for helping us understand the multifaceted topics of the groundfish fisheries, the observer program, and the people that rely on these fisheries for their livelihoods. In this regard we wish to especially thank Herman Savikko, Jennifer Mondragon, and Martin Loefflad. We thank Phil Mundy, Eric Volk, and Stephanie Moreland for several kinds of assistance, guidance, and council. Also, we thank Dave Carlile and Bill Templin of the Alaska Department of Fish and Game and Michele Masuda of the Ted Stevens Marine Research Institute for several suggestions that improved our presentation.

REFERENCES CITED

Anderson, E. C., R. S. Waples, and S. T. Kalinowski. 2008. An improved method for predicting the accuracy of genetic stock identification. Canadian Journal of Fisheries and Aquatic Sciences 65:1475–1486.

Anonymous. 2007. Report of the genetics workgroup. Pacific Salmon Commission. (download at http://www.psc.org/GSIWorkshop/2nd_Workshop_Sept2007/Day3/Genetics_Workgroup_Recommendations.pdf).

Anonymous. 2008. Bering Sea Chinook salmon bycatch management. Draft environmental impact statement/regulatory impact review/initial regulatory flexibility analysis. NOAA/NMFS and ADF&G. December 2008. 839p. (http://www.fakr.noaa.gov/sustainablefisheries/bycatch/salmon/deis1208.pdf).

Barns, A., M. Loefflad, and W. Karp. 2005. New fisheries monitoring and analysis division assumes the role of the North Pacific groundfish observer program. Alaska Fisheries Science Center, Quarterly Report (July-Aug-Sept 2005).

Beacham, T. D, K. D. Le, M. Wetklo. 2005. Summary of chum salmon stock identification research conducted by the Molecular Genetics Laboratory, Pacific Biological Station. (NPAFC Doc. 900). 6p. Department of Fisheries and Oceans, Pacific Biological Station, Nanaimo, B. C., Canada V9T 6N7.

Cahalan, J., and J. Ferdinand. 2008. Fisheries monitoring and analysis division redesigns observer sampling. Alaska Fisheries Science Center, Quarterly Report (Jan-Feb-Mar 2008).

Cochran, W. G., F. Mosteller, and J. W. Tukey. 1954. Principles of sampling. Journal of the American Statistical Association. 49:13–35 (originally appeared as Appendix G in *Statistical Problems with the Kinsey Report*).

Cochran, W. G. 1977. Sampling Techniques, 3rd ed. John Wiley and Sons. New York.

Debevec E., R. Gates, M. Masuda, J. Pella, J. Reynolds, and L. Seeb. 2000. SPAM (version 3.2): statistics program for analyzing mixtures. Journal of Heredity 91:509–510.

Edgington, E. S. 1980. Randomization tests. Marcel Dekker, New York.

Fisher, R. A. 1925. Statistical Methods for Research Workers, 1st ed. Oliver and Boyd. Edinburgh.

Fitzpatrick, S., and A. Scott. 1987. Quick simultaneous confidence intervals for multinomial proportions. Journal of the American Statistical Association 82:875–878.

Fournier, D., T. Beacham, B. Riddell, and C. Busack. 1984. Estimating stock composition in mixed stock fisheries using morphometric, meristic, and electrophoretic characteristics. Canadian Journal of Fisheries and Aquatic Sciences 41:400–408.

Garvin, M. R., and A. J. Gharrett. 2009. Developing DNA markers for the analysis of chum salmon bycatch in Alaskan trawl fisheries. Final Report submitted to Pollock Conservation Cooperative Research Center, School of Fisheries and Ocean Sciences, 245 O'Neill Building, University of Alaska Fairbanks, Fairbanks, AK, 34p.

Gharrett, A. J., S. A. Fuller, R. J. Riley, S. L. Hall, V. Brykov, and R. I. Wilmot. 2007. Developing DNA markers for the analysis of Bering Sea chum salmon. 2007 Arctic Yukon Kuskokwim Sustainable Salmon Initiative Project Final Product. August 2007, 70 p. + appendices.

Habicht, C., L. W. Seeb, K. Pronzati, and J. E. Seeb. 2006. SNPs provide in-season stock estimates for managing sockeye salmon in Bristol Bay: insights on stock-specific migration pathways and schooling off-shore. Abstract In SNP Workshop II: Applications of SNP genotyping in fisheries management. Alyeska Resort, Girdwood, Alaska, Sept 21-22, 2006. (http://www.genetics.cf.adfg.state.ak.us/workshop/AbstractFinal.pdf).

Hand, D. J. 1981. Discrimination and classification. John Wiley & Sons, New York, NY, 218p.

Hankin, D., J. Clark, R. Deriso, J. Garza, G. Morishima, B. Riddell, C. Schwarz, and J. Scott. 2005. Report of the expert panel on the future of the coded wire tag recovery program for Pacific salmon. An online version is available at www.psc.org/info_codedwiretagreview.htm.

Kalinowski, S. T. 2003. Genetic Mixture Analysis 1.0. Bozeman, MT: Department of Ecology, Montana State University. Available at http://www.montana.edu/kalinowski.

REFERENCES CITED (Continued)

Kalinowski, S. T. 2004. Genetic polymorphism and mixed-stock fisheries analysis. Canadian Journal of Fisheries and Aquatic Sciences 61:1075–1082.

Koljonen, M.-L., and R. Wilmot. 2005. Genetic analysis: allozymes. *In* Stock identification methods: applications in fishery science (S. Cadrin, K. Friedland, J. Waldman, editors), pp. 295–309, Elsevier Academic Press, New York, 719p.

Margoulas, A. 2005. Mitochondrial DNA. *In* Stock identification methods: applications in fishery science (S. Cadrin, K. Friedland, J. Waldman, editors), pp. 311–330, Elsevier Academic Press, New York, 719p.

Marine Resources Assessment Group Americas, Inc. 2000. Independent review of the North Pacific Groundfish Observer Program. Report to: Alaska Fisheries Science Center, National Marine Fisheries Service. (Unpublished)

Marine Resources Assessment Group Americas, Inc. 2003. Evaluation and analysis of current field sampling used in North Pacific groundfish fisheries. Report to: Alaska Fisheries Science Center, National Marine Fisheries Service. (Unpublished)

McKinnell, S., J. J. Pella, and M. L. Dahlberg. 1997. Population-specific aggregations of steelhead trout (*Oncorhynchus mykiss*) in the North Pacific Ocean. Canadian Journal of Fisheries and Aquatic Sciences 54:2368–2376.

Megrey, B. A. and V. G. Wespestad. 1990. Alaska groundfish resources: 10 years of management under the Magnuson Fishery Conservation and Management Act. North American Journal of Fisheries Management 10(2):125–143.

Millar, R. B. 1987. Maximum likelihood estimation of mixed stock fishery composition. Canadian Journal of Fisheries and Aquatic Sciences 44:583–590.

Miller, T. J. 2005. Estimation of catch parameter from a fishery observer program with multiple objectives. Ph.D. Thesis. University of Washington. Seattle, Washington.

Moran, P., M. Banks, T. Beacham, C. Garza, S. Narum, M. Powell, M. Campbell, L. Seeb, R. Wilmot, S. Young, B. Ardren, and J. Wenburg. 2005. Interlaboratory standardization of coast-wide Chinook salmon genetic data for international harvest management. Final report. A progress report from the Genetic Analysis of Pacific Salmonids (GAPS) consortium1 to the Chinook Technical Committee of the Pacific Salmon Commission, FY2004, FY2005.

Moran, P., M. Banks, T. Beacham, S. Narum, M. Campbell, L. Seeb, and S. Blankenship. 2006. An annual progress report from the Genetic Analysis of Pacific Salmonids (GAPS) consortium to the Chinook Technical Committee of the Pacific Salmon Commission, FY2006.

Myers, K.W. and D.E. Rogers. 1988. Stock origins of Chinook salmon in incidental catches by groundfish fisheries in the Eastern Bering Sea. North American Journal of Fishery Management 8(2):162–171.

Pella J., and G. Milner. 1987. Use of genetic marks in stock composition analysis. [*In*] Population genetics and fisheries management (N Ryman, F. Utter), pp. 247–276. Seattle, WA: University of Washington Press.

Pella, J., R. Rumbaugh, L. Simon, and M. Dahlberg. 1993. Incidental and illegal catches of Salmonids in North Pacific driftnet fisheries. Int. N. Pac. Fish. Comm. Bull. 53(III): 325-358.

Pella, J., M. Masuda, and S. Nelson. 1996. Search algorithms for computing stock composition of a mixture from traits of individuals by maximum likelihood. U.S. Department of Commerce, NOAA Tech. Memo, NMFS-AFSC-61.

Pella, J, and M. Masuda. 2001. Bayesian methods for analysis of stock mixtures from genetic characters. Fishery Bulletin 99(1):151–167.

Pella J, and M. Masuda. 2005. Classical discriminant analysis, classification of individuals, and source population composition of mixtures. *In* Stock identification methods: applications in fishery science (S. Cadrin, K. Friedland, J. Waldman, editors), pp. 517–552, Elsevier Academic Press, New York, 719p

REFERENCES CITED (Continued)

Sato, S., Y. Moongeun, S. Abe, and S. Urawa. 2007. Update of mitochondrial DNA baseline for stock identification of chum salmon. (NPAFC Doc. 1019) 26p. National Salmon Resources Center, Fisheries Research Agency, 2-2 Nakanoshima, Toyohira-ku, Sapporo 062-0922, Japan.

Smith, C. T., and L. W. Seeb. 2008. Number of alleles as a predictor of the relative assignment accuracy of short tandem repeat (STR) and single-nucleotide-polymorphism (SNP) baselines for chum Salmon. Transactions of the American Fisheries Society 137:751-762.

Seeb, J. E., S. Abe, S. Sato, S. Urawa, N. Varnavskaya, N. Klovatch, E. V. Farley, C. Guthrie, B. Templin, C. Habicht, J. M. Murphy, and L. W. Seeb. *in press*. The use of genetic stock identification to determine the distribution, migration, early marine survival, and relative stock abundance of sockeye, chum and Chinook salmon in the Bering Sea. 2008 BASIS SYMPOSIUM 54 3-1. Migration and Distribution of Salmon. NPAFC Bull. No. 5.

Seeb, J.E., C.E. Pascal, R. Ramakrishnan, and L.W. Seeb. 2009. SNP genotyping by the 5'-nuclease reaction: advances in high throughput genotyping with non-model organisms. A. Komar, editor. Methods in Molecular Biology, Single Nucleotide Polymorphisms, 2nd Edition. Humana Press (*chapter 18 in press*).

Seeb, L. W., A. Antonovich, M. Banks, T. Beacham, R. Bellinger, S. Blankenship, M. Campbell, N. Decovich, J. C. Garza, C. Guthrie, T. Lundrigan, P. Moran, S. Narum, J. Stephenson, J. Supernault, D. Teel, W. D. Templin, J. K. Wenburg, S. Young, and C. T. Smith. 2007. Development of a standardized DNA database for Chinook salmon. Fisheries 32(11):540–552.

Seeb, J. E., S. Abe, S. Sato, S. Urawa, N. Varnavskaya, N. Klovatch, E. V. Farley, C. Guthrie, B. Templin, C. Habicht, J. M. Murphy, and L. W. Seeb. *in press*. The use of genetic stock identification to determine the distribution, migration, early marine survival, and relative stock abundance of sockeye, chum and Chinook salmon in the Bering Sea. 2008 BASIS SYMPOSIUM 54 3-1. Migration and Distribution of Salmon. NPAFC Bull. No. 5.

Shaklee, J. B., T. D. Becham, L. Seeb, and B. A. White. 1999. Managing fisheries using genetic data: case studies from four species of Pacific salmon. Fisheries Research 43(1-3): 45-78.

Teel, D. J., P. A. Crane, C. M. Guthrie III, A. R. Marshall, D. M. Van Doornik, W. D. Templin, N. V. Varnavskaya, and L. W. Seeb. 1999. Comprehensive allozymes database discriminates Chinook salmon around the Pacific Rim. (NPAFC document 440) 25p. Alaska Department of Fish and Game, Division of Commercial Fisheries, 333 Raspberry Road, Anchorage Alaska USA 99518.

Templin, W. D., L. W. Seeb, J. Murphy, and J. E. Seeb. *in press*. High-Resolution Stock Identification for Migratory Studies of Chinook Salmon. 2008 BASIS SYMPOSIUM 54 3-1. Migration and Distribution of Salmon. NPAFC Bulletin No. 5.

Thompson, S.K. 1987. Sample size for estimating multinomial proportions. The American Statistician 41:42–46.

Thompson, S.K. 2002. Sampling, 2nd ed. Wiley Interscience. New York.

United States and Canada Yukon River Joint Technical Committee. 2007. Yukon River salmon 2006 season summary and 2007 outlook. Regional Information Report 3A07-01. Alaska Department of Fish and Game. Anchorage, Alaska.

Vϕlstad, J. H., W. Richkus, S. Gaurin, and R. Easton. 1997. Analytical and statistical review of procedures for collection and analysis of commercial fishery data used for management and assessment of groundfish stocks in the U.S. exclusive economic zone off Alaska. Report to: Alaska Fisheries Science Center, National Marine Fisheries Service. (Unpublished)

Wilmot, R. L., C. M. Kondzela, C. M. Guthrie, and M. M. Masuda. 1998. Genetic stock identification of chum salmon harvested incidentally in the 1994 and 1995 Bering Sea trawl fishery. North Pacific Anadromous Fish Commission Bulletin 1:285–299.

REFERENCES CITED (Continued)

Wilmot, R. L., C. M. Kondzela, C. M. Guthrie III, A. Moles, J. Pella, and M. Masuda. 2000. Origins of salmon seized from the F/V *Arctic Wind*. (NPAFC Doc.) Auke Bay Fisheries Laboratory, Alaska Fisheries Science Center, NMFS, NOAA, 11305 Glacier Highway, Juneau, AK 99801-8626.

Witherell, D. and C. Pautzke. 1997. A brief history of bycatch management measures for eastern Bering Sea groundfish fisheries. Marine Fisheries Review 59(4):15–22.

Witherell, D., D. Ackley, and C. Coon. 2002. Overview of salmon bycatch in Alaska groundfish fisheries. Alaska Fishery Research Bulletin 9(1):53–64.

Witherell, D. and D. Woodby. 2005. Application of marine protected areas for sustainable production and marine biodiversity off Alaska. Marine Fisheries Review (67):1–27.

Wood, C., S. McKinnell, T. Mulligan, and D. Fournier. 1987. Stock identification with the maximum-likelihood mixture model: sensitivity analysis and applications to complex problems. Canadian Journal of Fisheries and Aquatic Sciences 44:866–881.

Woodby, D., D. Carlile, S. Siddeek, F. Funk, J. H. Clark, and L. Hulbert. 2005. Commercial fisheries of Alaska. Alaska Department of Fish and Game Special Publication 05-09. Alaska Department of Fish and Game. Anchorage, Alaska.

APPENDICES

Appendix A.–Observer Forms and Information Recorded on Trawlers and at Plants

A total of seven different kinds of forms are completed by observers and these describe the information obtained:

1—Trip Form. All vessel observers must fill out this form, which captures position and calendar data at the trip level. This information includes the ports at the beginning and end of a trip, start and end dates and time, and time lost at sea.

2—Offload Form. All plant observers and all observers on catcher vessels delivering to a shore-based processor or a floating processor must fill out this form together. The form includes the cruise number, vessel permit number, vessel or plant name, observer's name, trip number, completion date, total delivery weight, total pollock weight, and the Alaska Department of Fish and Game fish ticket number.

3—Vessel Haul Form. All vessel observers must fill out this form, which captures geographic position at the haul level. An entry is required for every day at sea whether fished or not. All hauls retrieved must be recorded, and in the order of occurrence. Each haul has recorded the cruise number, vessel permit number, the observer's name, the vessel name, trip number, Community Development Quota group number if applicable, vessel type (catcher-processor, mothership, or catcher), gear problems during haul, date, time, latitude and longitude at beginning and end, average bottom depth and average gear depth during the deployment.

4—Observer Haul Form. All vessel observers must complete this form, which captures haul specific weight and other data. This form has a line corresponding to each line on the Vessel Haul Form to provide information about sampling or not sampling the hauls, and records the vessel's total catch weight estimate for the haul, any discard weight, the observer's total catch weight estimate and how estimated (bin volume, weighed entirely on observer scales or vessel's flow scale, or codend estimate), and delivery locations and dates.

5—Deck Form. All vessel observers must fill out this waterproof form, which contains species composition and other data specific to species composition sampling. Information for each sampled haul is recorded including date of retrieval, cruise number, vessel permit number, haul or offload number, sample and subsample numbers and weights, and a list of species present (by sex for Tanner crab, king crab, and salmon) with corresponding counts of the individuals and their combined weights. As of January 2008, observers are to collect and individually record at least three collections for species composition from each sampled haul. Completion of paper versions called the Species Composition Form and the Length and Specimen Form are also necessary if faxing is the mode of communication for the vessel.

6—Marine Mammal Interaction and Specimen Form. When marine mammal interactions occur, observers on vessels faxing data must fill out this paper form, which captures information about each event. Observers on vessels communicating with the ATLAS system transmit the same information electronically.

7—Species Identification Form. All observers, whether at a plant or on a vessel, must fill out this form for each new species they see. The form captures characters of the species and is used to verify the observer's species identification.

Miller (2005) developed the general framework and formulas whereby the catch of any species, such as Chinook salmon bycatch, could be estimated from observer data. Miller's work was completed by 2005, and the observer program and regulations have changed since then to allow improved estimation. Miller provided design-based as well as model-based methods of estimation. Design-based estimators require no assumptions about the populations being sampled, but they do require assumptions about the selection of the sample. Miller's model-based estimators do require assumptions about the nature of the bycatch, but these appear so plausible that any disadvantage is minor. Miller developed the model-based estimators for limited data situations that prevented use of design-based estimators. The limited data situations would seem more frequent in the earlier years of the observer program than nowadays. For example, observers are now required to obtain at least three species composition samples from a sampled haul (see p. 5-5 of 2008 Observer Sampling Manual: http://www.afsc.noaa.gov/ FMA/ document.htm) and these data are recorded separately for the samples in the Alaska Fisheries Science Center database (Cahalan and Ferdinand 2008). Previously, fewer species composition samples were taken and the data were aggregated over samples before entry into the database. In addition, Miller viewed the numbers of trips made by midsized vessels as unknown. We believe that such information is available today and can be obtained from the fish tickets at offloading required by the Alaska Department of Fish and Game. Therefore, Miller's approach to dealing with this uncertainty in numbers of trips is omitted.

Design-based estimation within a haul

All, or nearly all, hauls by a catcher vessel should be sampled because at most a few hauls are made per day. However, if all the hauls cannot be sampled, the observer is directed to introduce randomness in the selection of hauls to be sampled. True random sampling is not possible because of the physical limits to which the observer could be pressed and because the number of hauls is not known before sampling begins. As Miller notes, treating the hauls as though randomly sampled will yield appropriate inferences because the haul contents are not related to their inclusion for sampling. The species composition samples are to be drawn either systematically or randomly from the chosen sample frame. The frame is either temporal or spatial. A temporal frame is available if the catch passes a sampling location on the vessel slowly enough so that portions can be removed for examination, or a spatial frame is available if the catch is brought onto the vessel at a sampling location where it can be partitioned into portions that can be removed for examination. A measure of size, either weight or volume, of each sample is obtained, and the corresponding size of the unsorted haul catch is recorded. As a result of Amendment 80, nearly all catcher/processor vessels in the Bering Sea now have motion-compensated scales available to make the size, i.e., weight, measurement of samples more accurate. The estimator for total number of Chinook salmon in the *k*th haul is (Miller's Equation 3.10),

$$\hat{N}_k = \frac{V_k}{v_k m_k} \sum_{i=1}^{m_k} N_{ki} = \frac{M_k}{m_k} \sum_{i=1}^{m_k} N_{ki} \,, \qquad (1)$$

and the estimate for the variance is (Miller's Equation 3.11),

$$\hat{V}\left(\hat{N}_k\right) = M_k\left(\frac{M_k}{m_k} - 1\right)\frac{\sum_{i=1}^{m_k}\left(N_{ki} - \hat{\bar{N}}\right)^2}{m_k - 1}. \qquad (2)$$

Here N_{ki} denotes the number of Chinook salmon counted in the ith sample, V_k equals the total size of the haul, v_k is the size of each equal-sized sample, a total of m_k sampling units are taken from the haul which contains a total of M_k sampling units, and $\hat{\bar{N}} = \sum_{i=1}^{m_k} N_{ki} \Big/ m_k$.

Miller also adjusts the variance estimator upward for the situation where the size measurements of the sampling units are subject to error, but the correction (Miller's Equation 3.13) requires experimentally determined values for the unknown measurement error variances for every measurement method.

Model-based estimation within a haul

Miller develops a ratio estimator of the bycatch of a haul that may prove useful when actual sampling on a vessel carrying an observer is incomplete. Furthermore, unobserved midsized vessels may provide information at delivery of catch to the processor by which to estimate haul or trip bycatch. Such model-based estimates may be useful for comparison with design-based estimates if dumping is suspected. The accuracy of design-based estimates depends on the assumption that bycatch of a vessel is not affected by the presence of an observer. The generic model (actually, Miller described two models for this problem and we limit discussion to that which seems most realistic) for bycatch of the kth haul is

$$\begin{aligned} E_M\left(N_k\right) &= x_k \phi_{N,t} \\ V_M\left(N_k\right) &= x_k \sigma_{N,t} \\ Cov_M\left(N_j, N_k\right) &= 0 \quad \text{for } j \neq k, \end{aligned} \qquad (3)$$

where $E_M\left(\bullet\right)$ denotes the averaging process of the argument with respect to the model, x_k is the known value of a covariate measured on the haul such as pollock weight, $\phi_{N,t}$ is a trip-specific slope parameter such as the bycatch rate of Chinook salmon per ton of pollock on a trip t, and $\sigma_{N,t}$ is a trip-specific measure of the change in variation of bycatch as the magnitude of the covariate changes. The weight of pollock catch may be the best covariate because at least its trip value will usually be available. Miller derives the optimal estimator for $\phi_{N,t}$,

$$\hat{\phi}_{N,t} = \sum_{k=1}^{g_t}\sum_{i=1}^{m_k} N_{ki} \Big/ \sum_{k=1}^{g_t} x_k , \qquad (4)$$

where the numerator equals the total count of Chinook salmon among the species composition sample collections (number of species composition sample collections drawn from the kth haul is m_k) of the g_t sampled hauls of the trip, and the denominator is the corresponding sum of the covariate for the species composition sample collections. The estimator of the total bycatch in the kth haul is

$$\hat{N}_k = \sum_{i=1}^{m_k} N_{ki} + \hat{\phi}_{N,t} x'_k , \tag{5}$$

where x'_k denotes the covariate total over the unsampled portion of the kth haul. Notice that the estimator comprises two terms, the first equal to the total count of Chinook salmon bycatch in the m_k species composition samples from the haul, and the second equal to the corresponding estimated bycatch in the unobserved portion of the haul. The unbiased estimator of the variance for the bycatch estimator, $\hat{N}_k$, is (Miller's Equation 3.36 and nearby)

$$\hat{V}_M\left(\hat{N}_k\right) = \left(\frac{x'_k}{\sum_{k=1}^{g_t} x_k} + 1 \right) \cdot x'_k \cdot \hat{\sigma}^2_{N,t} \tag{6}$$

where $$\hat{\sigma}^2_{N,t} = \frac{\left(\sum_{k=1}^{g_t} x_k\right)}{\left(\sum_{k=1}^{g_t} x_k\right)^2 - \sum_{k=1}^{g_t} x_k^2} \sum_{k=1}^{g_t} \left(\hat{N}_k - \hat{\phi}_{N,t} x_k\right)^2 .$$

Design-based estimation for a trip

A catcher vessel with an observer on board may make one or more hauls before delivery to a processing vessel. A catcher-processor with an observer on board continues to make hauls until filled. The total bycatch of Chinook salmon is estimated from the individual hauls. Suppose that the observer has sampled g_t hauls for species composition from the total of G_t hauls of the trip. Miller treats the sampled hauls as if randomly sampled from the total hauls of the trip (although the captain of the midsized vessels chooses the trips to be observed) and uses the expanded sum of the within-haul estimates for the total (Miller's Equation 5.1),

$$\hat{N}_t = \frac{G_t}{g_t} \sum_{k=1}^{g_t} \hat{N}_k , \tag{7}$$

and the estimated variance is (Miller's Equation 5.2),

$$\hat{V}\left(\hat{N}_t\right) = G_t \left(\frac{G_t}{g_t} - 1 \right) \tilde{S}_t^2 + \frac{G_t}{g_t} \sum_{k=1}^{g_t} \hat{V}\left(\hat{N}_k\right) \tag{8}$$

Here $\tilde{S}_t^2 = \dfrac{\sum_{k=1}^{g_t} \left(\hat{N}_k - \hat{\bar{N}}_t\right)^2}{g_t - 1}$, $\hat{\bar{N}}_t = \sum_{k=1}^{g_t} \hat{N}_k \Big/ g_t$, and the terms $\hat{N}_k$ and $\hat{V}\left(\hat{N}_k\right)$ refer to the design-based haul estimators defined earlier.

In the special case when only a single haul is made in the trip, if the observer sampled the haul, the trip bycatch and its variance are simply estimated from the within-haul formulas,

$$\hat{N}_t = \hat{N}_{k=1} \text{ and } \hat{V}\left(\hat{N}_{k=1}\right).$$

If the observer did not sample the haul, then Miller uses model-based methods for the bycatch estimation (see Equations 18–20 below).

Model-based estimation for a trip

The estimate of the trip bycatch from the model is

$$\hat{N}_t = \sum_{k=1}^{g_t}\sum_{i=1}^{m_k} N_{ki} + \hat{\phi}_{N,t}\left(\sum_{k=1}^{g_t} x'_k + \sum_{l=1}^{g'_t} X_l\right) = N_{++,t} + \hat{\phi}_{N,t} x'_t \,, \tag{9}$$

where x'_k is the covariate total for the unsampled portion of the *k*th haul, g'_t is the total number of unsampled hauls in the *t*th trip, $N_{++,t}$ is the total number of Chinook (or chum) salmon counted in species composition samples from the hauls during trip *t*, x'_t is the covariate total for unsampled hauls and unsampled portions of sampled hauls of trip *t*, $\hat{\phi}_{N,t}$ is the estimated slope coefficient at Equation 4, and X_l is the covariate total for the *l*th unsampled haul. Miller provides the prediction error variance estimator for trip bycatch, $\hat{N}_t$,

$$V_M\left(\hat{N}_t\right) = X_t \frac{x'_t}{x_t} \hat{\sigma}^2_{N,t} \,, \tag{10}$$

where X_t is the covariate total for the entire trip *t*, x_t is the covariate total for sampled portions of hauls for the trip, x'_t is the covariate total for the unsampled portions of hauls for the trip, and

$$\hat{\sigma}^2_{N,t} = \frac{x_t}{x_t^2 - \sum_{k=1}^{g_t} x_k^2} \sum_{k=1}^{g_t}\left(N_{k+,t} - \hat{\phi}_{N,t} x_k\right)^2 \,, \tag{11}$$

where $N_{k+,t} = \sum_{i=1}^{g_t} N_{ki}$ is the total count of Chinook (or chum) salmon in the species composition samples from the *k*th haul.

Design-based estimation of domain bycatch of a vessel trip

Managers will need to know the bycatch for time and area domains for a fleet of vessels. However, any of these vessels may have made hauls during a trip within a domain as well as outside of the domain, Miller begins with estimating the *d*th domain total bycatch for a particular vessel trip *t* as the sum (Miller's Equation 5.40),

$$\hat{N}_{dt} = \frac{G_t}{g_t} \sum_{k=1}^{g_{dt}} \hat{N}_k \,, \tag{12}$$

where G_t is the number of hauls during the trip, g_t is number of sampled hauls in the trip, g_{dt} is the number of the sampled hauls that occurred in domain *d* during the trip, and $\hat{N}_k$ is the estimator (see Equation 1 or 5) of the bycatch in the *k*th sampled haul. The variance of this estimator is (Miller's Equation 5.41),

$$\hat{V}\left(\hat{N}_{dt}\right) = G_t\left(\frac{G_t}{g_t} - 1\right)\left[\frac{\sum_{k=1}^{g_{dt}} \hat{N}_k^2 - \left(\sum_{k=1}^{g_{dt}} \hat{N}_k\right)^2 \Big/ g_t}{g_t - 1}\right] + \frac{G_t}{g_t}\sum_{k=1}^{g_{dt}} \hat{V}\left(\hat{N}_k\right). \tag{13}$$

Model-based estimation for domain *d* bycatch during vessel trip *t*

When the covariate total, e.g., pollock catch, for the domain *d* during trip *t* is known, Miller uses a model with domain-specific parameters,

$$\begin{aligned} E_M\left(N_k\right) &= x_k \phi_{N,dt} \\ V_M\left(N_k\right) &= x_k \sigma_{N,dt}^2 \\ Cov_M\left(N_k, N_l\right) &= 0 \quad \text{for } k \neq l, \end{aligned} \tag{14}$$

where N_k is the count of Chinook (or chum) salmon in the sampled portion of the *k*th haul, x_k is the corresponding covariate total in that haul, and $\phi_{N,dt}$ and $\sigma_{N,t}$ are the domain slope and change in variation parameters. The estimator of the bycatch in the domain during a trip sums the observed counts in sampled hauls and estimated counts for any unsampled hauls or portions of hauls as (Miller's Equation 5.38),

$$\hat{N}_{dt} = \sum_{k=1}^{g_{dt}} N_k + \hat{\phi}_{N,dt}\left(\sum_{k=1}^{g_{dt}} x'_k + \sum_{l=1}^{g'_{dt}} X_l\right) = N_{dt} + \hat{\phi}_{N,dt} \cdot x'_{dt}, \tag{15}$$

where g'_{dt} is the number of unsampled hauls in the *d*th domain during the *t*th trip, g_{dt} is the number of hauls sampled in the *d*th domain during the *t*th trip, x'_k is the covariate total for the unsampled portion of the *k*th sampled haul, X_l is the covariate total for the *l*th unsampled haul, N_{dt} is the count of Chinook (or chum) salmon in sampled portions of the hauls, and x'_{dt} is the covariate total for unsampled portions of sampled hauls and unsampled hauls. The unbiased estimator for the slope parameter is

$$\hat{\phi}_{dt} = \sum_{k=1}^{g_{dt}} N_k \Big/ \sum_{k=1}^{g_{dt}} x_k. \tag{16}$$

Finally, the variance estimator for the bycatch in the domain during trip *t* is (Miller's Equation 5.38) is

$$\hat{V}_M\left(\hat{N}_{dt}\right) = \frac{x'_{dt}}{x_{dt}} X_{dt} \hat{\sigma}_{N,dt}^2, \tag{17}$$

where
$$\hat{\sigma}_{N,dt}^2 = \frac{x_{dt}}{x_{dt}^2 - \sum_{k=1}^{g_{dt}} x_k^2} \sum_{k=1}^{g_{dt}} \left(N_k - \hat{\phi}_{N,dt} x_k\right)^2,$$

$x_{dt} = \sum_{k=1}^{g_{dt}} x_k$, $X_{dt} = \sum_{k=1}^{G_{dt}} X_k$, and G_{dt} is the number of hauls made in the *d*th domain during the *t*th vessel trip.

Model-based estimation for undersampled trips

Some trips with an observer present may have had an insufficient number of hauls sampled for variance estimation. Useful information about the number of hauls made during the trip, the location and date of each haul, and the covariate value for each haul may be available. Miller uses a model closely related to that described earlier at Equation 14 except that the coefficients are now domain-specific rather than trip-specific,

$$\begin{aligned} E_M\left(N_k\right) &= x_k \phi_{N,m} \\ V_M\left(N_k\right) &= x_k \sigma^2_{N,m} \\ Cov_M\left(N_k, N_l\right) &= 0 \quad \text{for } k \neq l \end{aligned} \tag{18}$$

The predictor for bycatch in the undersampled trip is

$$\hat{N}_{mt} = X_{mt} \hat{\phi}_{N,m}\,, \tag{19}$$

where $\hat{\phi}_{N,m} = \sum_{k=1}^{g_m} N_k \Big/ \sum_{k=1}^{g_m} x_k$, X_{mt} is the covariate total for the hauls made on the undersampled trip t in the mth domain, and g_m hauls were sampled in the domain. The variance estimator is

$$\hat{V}\left(\hat{N}_{mt}\right) = X_{mt}\left(1 + \frac{X_{mt}}{x_m}\right)\hat{\sigma}^2_{N,m}\,, \tag{20}$$

where

$$\hat{\sigma}^2_{N,m} = \frac{x_m}{x_m^2 - \sum_{k=1}^{g_m} x_k^2} \sum_{k=1}^{g_m}\left(N_k - \hat{\phi}_{N,m} x_k\right)^2\,, \tag{21}$$

and x_m is the covariate total for the sampled portions of hauls made in the mth domain.

Design-based estimation for bycatch by a vessel over a number of trips

Miller assumes simple random sampling of c_v trips from a total of C_v trips by a particular vessel v. The estimator for total bycatch of Chinook (or chum) salmon is the sum of the trip bycatches (Miller's Equation 5.75),

$$\hat{N}_v = \frac{C_v}{c_v} \sum_{t=1}^{c_v} \hat{N}_t \tag{22}$$

and the variance estimator (Miller's Equation 5.76) is

$$\hat{V}\left(\hat{N}_v\right) = C_v\left(\frac{C_v}{c_v} - 1\right)\tilde{S}^2_{N,v} + \frac{C_v}{c_v}\left[\sum_{t=1}^{c_v}\hat{V}\left(\hat{N}_t\right) + \frac{C_v - 1}{c_v - 1}\sum_{t=1}^{c_u}\sum_{\substack{u=1 \\ u \neq t}}^{c_u} C\hat{o}v\left(\hat{N}_t, \hat{N}_u\right)\right], \tag{23}$$

where $\tilde{S}^2_{N,v} = \dfrac{\sum\left(\hat{N}_t - \hat{\bar{N}}_v\right)^2}{c_v - 1}$, $\hat{\bar{N}}_v = \hat{N}_v / C_v$, and $\hat{N}_t$ is an appropriate estimator of the bycatch for the *t*th trip.

The covariances of Equation 23 are all zero unless a vessel has two or more undersampled trips. Presumably that event will be so infrequent as to be ignorable.

Appendix C.–The General Mixture Problem and its Solution.

The source populations potentially include individuals of all the possible genotypes that occur among all populations in a mixture. Hence none of the populations can be excluded as the source of an individual. However, the relative frequency of any genotype generally differs among the populations, and so in this sense the plausibility that the individual came from any population varies: a population in which the genotype of an individual has a higher relative frequency than in another population is a more plausible candidate as the source provided that both populations were considered equally probable candidates prior to observing the genotype. Suppose that a random sample of M individuals is drawn from a large mixture composed of c populations with proportions $\mathbf{p} = (p_1,\ldots,p_c)$, $0 \le p_i \le 1$, $i = 1,\ldots,c$, and $\sum_{i=1}^{c} p_i = 1$. Alternatively, the sample could have been assembled as a composite of randomly selected individuals from the c populations by some unknown procedure. In either case, the unknown population composition of the mixture sample will be assessed from the genotypes of its members and used to interpret the underlying population mixture. The genotypes of the mixture individuals are determined in the laboratory at each of L independent loci with the hth locus having a total of J_h different alleles among the possible populations. The collection of genotypes is denoted as $\mathbf{X} = (\mathbf{X}_1,\ldots,\mathbf{X}_m)$. The genotype of the mth individual is $\mathbf{X}_m = (\mathbf{X}_{m1},\ldots,\mathbf{X}_{mL})$, where $\mathbf{X}_{mh} = (x_{mh1},\ldots,x_{mhJ_h})$ is the vector of allele counts (entries are 0, 1, or 2 for diploid animals such as fish) for an individual m at the hth locus. (Notice also that the total allele counts for a diploid individual at any locus h is $x_{mh+} = \sum_{j=1}^{J_h} x_{mhj} = 2$.) Hardy-Weinberg linkage equilibrium conditions (HWLE) are quite plausible for large panmictic populations that do not interbreed, and these conditions are assumed to hold for the salmon populations of this study. (Notice that if microsatellites—less likely—or the large numbers of SNPs—more likely—are in linkage disequilibrium, the computations that follow do not apply.) Therefore, the relative frequency of the genotype $\mathbf{X}_m$ in the ith population can be written as

$$f\left(\mathbf{X}_m;\mathbf{Q}_i\right) = \prod_{h=1}^{L}\left[2^{1-\delta_h(\mathbf{X}_m)}\prod_{j=1}^{J_h} q_{ihj}^{x_{mhj}}\right].$$

Here $\delta_h\left(\mathbf{X}_m\right)$ is the indicator function for the homozygous condition at locus h (equals 1 if the mth individual has $x_{mhj} = 2$ for some j, $j = 1,\ldots,J_h$, and equals 0 otherwise). Also, the genetic parameters for the ith stock are denoted by $\mathbf{Q}_i = (q_{ihj})$, $h = 1,\ldots,L$; $j = 1,\ldots,J_h$ and refer to the allele relative frequencies at the L loci defining the multilocus genotypes of individuals ($0 \le q_{hj} \le 1$, $\sum_{j=1}^{J_h} q_{hj} = 1$).

The relative frequency of the genotype for the mth individual in the mixture is a weighted average of its relative frequencies in the separate populations, $\sum_{i=1}^{c} p_i f\left(\mathbf{X}_m;\mathbf{Q}_i\right)$. Further the probability of the sample of mixture genotypes $\mathbf{X}$ is

$$\text{Prob}\left(\mathbf{X}_1,\ldots,\mathbf{X}_M\right) \propto \prod_{m=1}^{M}\left(\sum_{i=1}^{c} p_i f\left(\mathbf{X}_m;\mathbf{Q_i}\right)\right). \qquad (1)$$

Equation 1 is the basic genetic mixture model by which to estimate the population source composition of the catch. As written, the unknowns, **p** and **Q**, are viewed as constants, the genotypes, **X,** are treated as random variables, and likelihood methods are used for estimation. When Bayesian methods are introduced later, the view is reversed: the unknowns become random variables and the genotypes are the only constants. Both likelihood and Bayesian methods of genetic population mixture analysis are commonly used. Likelihood methods are simpler to use, but somewhat less efficient in use of information in the mixture sample. In addition, for some genetic systems in which allele relative frequencies are very small among populations, e.g., many microsatellites, an incongruous incorporation of Bayesian methods within the likelihood algorithms has been necessary. The full Bayesian method overcomes these issues.

Likelihood estimation method

The conditional maximum likelihood method is commonly used for population mixture analysis. The approach breaks the estimation problem into two steps: first, the genetic parameters are estimated from baseline samples taken from the separate populations, and second, the mixture proportions are estimated treating the genetic parameters as known and maximizing the probability function at Equation 1 with respect to the unknown **p**. Several algorithms have been proposed to estimate the unknowns by the conditional maximum likelihood method, e.g., see Fournier et al. (1984), Millar (1987), and Pella and Milner (1987). For example, the maximizing value of **p** given the genotypes can be obtained by iteratively (t = 1, 2,…, T) evaluating the equation system,

$$p_1^{(t)} = \frac{1}{M}\sum_{m=1}^{M}\frac{p_1^{(t-1)} f(\mathbf{X}_m;\mathbf{Q}_1)}{\sum_{i=1}^{c} p_i^{(t-1)} f(\mathbf{X}_m;\mathbf{Q}_i)} = \frac{1}{M}\sum_{m=1}^{M} P^{(t-1)}(1\,|\,\mathbf{X}_m)$$

$$\vdots$$

$$p_c^{(t)} = \frac{1}{M}\sum_{m=1}^{M}\frac{p_c^{(t-1)} f(\mathbf{X}_m;\mathbf{Q}_c)}{\sum_{i=1}^{c} p_i^{(t-1)} f(\mathbf{X}_m;\mathbf{Q}_i)} = \frac{1}{M}\sum_{m=1}^{M} P^{(t-1)}(c\,|\,\mathbf{X}_m)\ .$$

Here the expressions for posterior source probabilities at the iterations for each of the mixture individuals, m = 1,…,M, are introduced as

$$P^{(t-1)}\left(k\,|\,\mathbf{X}_m\right) = \frac{p_k^{(t-1)} f\left(\mathbf{X}_m;\mathbf{Q}_k\right)}{\sum_{i=1}^{c} p_i^{(t-1)} f\left(\mathbf{X}_m;\mathbf{Q}_i\right)} \quad ,k = 1,\ldots,c\,.$$

An arbitrary value for $\mathbf{p}^{(0)}$ with positive components of unit sum can be substituted into the right hand side of the equation system to start the iteration. The composition for $\mathbf{p}^{(1)}$ results on the left hand side, and it is substituted into the right hand side again. The process is continued to convergence, which can be judged to have occurred by various criteria. Typically convergence is assumed when the changes between successive estimates of **p** become arbitrarily small. This

method of computing the conditional maximum likelihood estimate $\hat{\mathbf{p}}$ is called the EM algorithm. Although the conditional MLE for $\mathbf{p}$ by the EM algorithm may not be as fast to compute as by other algorithms (Pella et al. 1996), it has the advantages of simplicity and the guarantee to converge with monotonic increase in likelihood function values during the search. Another advantage is the simple and intuitive interpretation of what is being done to compute the conditional MLE. Notice from the equation system that at convergence each individual has a unit value that is divided up among the possible stocks in proportion to its posterior source probabilities, and that the stock composition estimate is the arithmetic average of these posterior source probabilities over individuals of the sample. Therefore, instead of assigning the entire individual to a stock as though one knew its source, the individual is assigned fractionally in proportion to measures of our belief of its sources.

The estimation of the allele relative frequencies $\mathbf{Q}$ among loci and stocks has become more complex as choice in genetic marks changed. Initially, allozymes were available and these loci typically had only a few alleles per locus. Therefore, maximum likelihood estimation of the allele frequencies was based on the standard multinomial sampling model. Random samples of fish from escapements, called baseline samples, were assayed for their genotypes, *i.e.*, the pairs of alleles per fish. The maximum likelihood estimates of the $\mathbf{Q}_i$ are the observed allele compositions of the combined alleles of the baseline sample from each stock and locus. More recently, microsatellite loci were used for which the number of different alleles among stocks could be large (as many as 50 or even more) and many of these alleles were rare or in low relative frequency in stocks. The result was that sampling zeros were presumably common in the baseline samples and this became very problematic to stock composition estimation. If maximum likelihood was used to estimate the allele relative frequencies from the baseline samples, stocks whose baseline samples had sampling zeros were necessarily eliminated as potential sources for mixture individuals having the corresponding alleles in their genotypes. Other stocks without sampling zeros became the candidates, and with far greater posterior source probabilities than should have occurred. Because certainty about the absence of a rare allele is not possible with limited sample size from large populations, stock composition estimation from microsatellites performs better if the possibility is maintained that any baseline stock could be the source of any mixture individual. This possibility is achieved through use of Bayesian methods that in the present context provides a probability distribution for allele relative frequencies of $\mathbf{Q}$ rather than a point estimate and associated measure of variation. Under this Bayesian probability distribution, every allele is potentially present in every stock. Were a point estimate needed, the location parameter, e.g., mean or median, of the Bayes distributions of components of $\mathbf{Q}$ should suffice.

The recommended method of estimating the sampling variation in the conditional maximum likelihood estimate, $\hat{\mathbf{p}}$, is by bootstrap resampling. The reason that asymptotic methods are not useful for this purpose is they are inaccurate and overestimate the uncertainty. Asymptotic methods are inaccurate because they depend on an assumption, nearly always violated, that the distributions of composition estimates do not encounter the boundaries for the unknown proportions ($0 \le p_i \le 1,\ i = 1,\ldots,c,\ \sum_{i=1}^{c} p_i = 1$). In the bootstrap method, the mixture and baseline samples are sampled with replacement to generate random analogs of the same size. The allele relative frequencies $\mathbf{Q}$ are estimated by either maximum likelihood from the bootstrap baseline samples (appropriate if none of the alleles are rare), or by a single draw from the Bayes

distribution of **Q** (appropriate when some alleles are rare). Then the conditional maximum likelihood estimate of **p** is computed for the bootstrap mixture sample using the EM algorithm, for example. A large number of repetitions of this process (e.g., 1,000 times) generates an empirical bootstrap distribution from which the mean and lower $\alpha/2 \cdot 100$-*th* percentile and upper $(1-\alpha/2) \cdot 100$-*th* percentile provide a point estimate and symmetric $(1-\alpha) \cdot 100\%$ confidence bounds.

Software to perform the conditional maximum likelihood estimation method and bootstrap resampling of baseline and mixture samples is freely available from two sources. The earliest software is the program called Statistical Package for Analyzing Mixtures (SPAM) (Debevec et al. 2000). SPAM originally did not include the option during bootstrap resampling of using the Bayesian posterior for baseline allele relative frequencies. Meanwhile, Kalinowski (2003) developed the program called Genetic Mixture Analysis (GMA) that did. Later, SPAM was updated to include the option (Alaska Department of Fish and Game 2003).

Bayesian estimation method

The Bayesian approach has been extended to include estimation of both the stock composition **p** and the allele relative frequencies **Q** (Pella and Masuda 2001). The combination of a Bayesian analysis for **Q** and a maximum likelihood approach for **p** described above is a peculiarity in that frequentist and Bayesian statisticians view any estimation problem as mirror images. Under Bayesian methods, data are considered fixed and unknown parameters are considered random variables. Under frequentist methods such as maximum likelihood, data are considered random and parameters are fixed. Although the two schools have long debated the validity of their approaches, the more recent Bayesian methods have gained favor by more exact modeling of complex problems made possible by availability of greater computing power. The premise of the Bayesian method for estimation of a collection of unknowns, say $\Theta = (\mathbf{p}, \mathbf{Q})$, is that information exists about Θ before a sample is drawn and data **Y** become available. The information is provided in the form of a prior probability distribution, in the present problem by $\pi(\mathbf{p}, \mathbf{Q}) = \pi(\mathbf{p})\pi(\mathbf{Q})$. That is, the prior information about the stock proportions, **p**, and genetic parameters, **Q**, is assumed to be statistically independent so that their joint prior probability distribution equals the product of their separate prior distributions. Uninformative priors are chosen for both **p** and **Q** so that the data "do the talking." The information in the data **Y** is combined with that of the prior by integration of the product of the prior, $\pi(\mathbf{p}, \mathbf{Q})$, and the likelihood of the data, $\pi(\mathbf{Y} \mid \mathbf{p}, \mathbf{Q}) \propto \prod_{m=1}^{M} \left[\sum_{i=1}^{c} p_i f(\mathbf{X}_m \mid \mathbf{Q}_i) \right]$, to produce the posterior distribution, $\pi(\mathbf{p}, \mathbf{Q} \mid \mathbf{Y}) \propto \int \pi(\mathbf{p}, \mathbf{Q})) \cdot \pi(\mathbf{Y} \mid \mathbf{p}, \mathbf{Q}) d\mathbf{p} d\mathbf{Q}$. The posterior distribution summarizes the knowledge and uncertainty about the unknowns. Pella and Masuda (2001) chose so-called conjugate priors for both **p** and **Q** so that the Bayes posterior distribution for the unknowns has an explicit solution. The prior for **p** is the Dirichlet density function, which is defined on the stock composition simplex, $S(\mathbf{p}) = \left\{ \mathbf{p} : 0 < p_i < 1, \sum_{i=1}^{c} p_i = 1 \right\}$. The density function is

parameterized by $\alpha = (\alpha_1,\ldots,\alpha_c)$ and is $\pi(\mathbf{p} \mid \alpha) = D(\mathbf{p} \mid \alpha) = \dfrac{\Gamma\left(\sum_{i=1}^{c} \alpha_i\right)}{\prod_{i=1}^{c} \Gamma(\alpha_i)} \prod_{i=1}^{c} p_i^{\alpha_i - 1}$. Pella and Masuda (2001) set the $\alpha_i = 1/c$, $i = 1,\ldots,c$, which has the desired effect of providing low information about **p**: this prior information is equivalent to adding a single individual to the mixture sample and specifies that the prior contributions from the source stocks to the mixture are equal. If the source identities of the M individuals in the random multinomial mixture sample were available and the array of stock counts is denoted by $\mathbf{Z} = (z_1,\ldots,z_c)$, where $\sum_{i=1}^{c} z_i = M$, the posterior distribution is the Dirichlet density function, $\pi(\mathbf{p} \mid \mathbf{Z}) = D\left(\mathbf{p} \mid \mathbf{Z} + \frac{1}{c} \cdot \mathbf{1}\right) = D\left(\mathbf{p} \mid \mathbf{Z} + \frac{1}{c} \cdot (1,\ldots,1)\right)$. The means, variances, and covariances of this distribution are given by

$E(p_i \mid \mathbf{Z}) = (z_i + c^{-1})/(M + 1)$, $i = 1,\ldots,c$

$\text{var}\,(p_i \mid \mathbf{Z}) = [(z_i + c^{-1})(M + 1 - (z_i + c^{-1}))]/[(M + 1)^2(M + 2)]$, $i = 1,\ldots,c$

$\text{cov}\,(p_i, p_j \mid \mathbf{Z}) = [(z_i + c^{-1})(z_i + c^{-1})]/[(M + 1)^2(M + 2)]$, $i, j = 1,\ldots,c$

As M becomes large, these values for the posterior distribution of **p** agree closely with the MLE estimates of the corresponding values if **Z** is treated as a simple multinomial random variable, so that the Bayesian posterior distribution will be a reasonable description of knowledge and uncertainty for both Bayesian and frequentist statisticians. Of course, the actual counts of individuals by source stock are unknown and what the Bayesian method does to accommodate this uncertainty is described shortly.

The remaining unknowns are the allele relative frequencies **Q** at H different loci among the c stocks. Two information sources about **Q** are available: the baseline samples, and the mixture sample. In contrast to the conditional maximum likelihood method, which uses only the baseline samples for estimating **Q**, the Bayesian method extracts the information about **Q** from both. First, a separate Bayesian analysis is performed with the baseline samples to develop the baseline posterior distribution for **Q**. Second, the baseline posterior distribution for **Q** becomes the mixture prior for **Q** to be updated during the mixture sample analysis. The baseline sample for the *h-th* locus from the *i-th* stock is viewed as a random draw from the multinomial probability function with J_h different alleles possible at the locus. In practice, the value for J_h is the number of different alleles at the locus observed among the baseline stocks. Again, the Dirichlet prior probability density has been used to describe knowledge and uncertainty in **Q**. Two specifications for the parameters of the Dirichlet prior have been used. The most straightforward specification (Rannala and Mountain 1997) is the analog to that described above for stock composition **p**, but now applied to each of the $c \cdot H$ unknown allele relative frequency arrays denoted by $\mathbf{q}_{ih} = (q_{ih1},\ldots,q_{ihJ_h})$, $i = 1,\ldots,c$, $h = 1,\ldots,H$ (see Kalinowski 2003). To distinguish these prior parameters from those used for **p**, they will be denoted by $\beta_{ih} = (\beta_{ih1},\ldots,\beta_{ihJ_h})$, where $\beta_{ihj} = J_h^{-1}$ if the *h-th* locus has J_h alleles among the baseline stocks. The other specification by Pella and Masuda (2001) chooses the prior parameters for a locus to be proportional to a baseline center of allele relative frequencies. The baseline center is the unweighted arithmetic average of

the observed allele relative frequencies at the locus among the stocks in the baseline. The value for the constant of proportionality of each locus is chosen to minimize the sum of squared deviations between the observed allele relative frequencies and their posterior mean. With this definition for the prior parameters, the prior mean equals the baseline center, and the posterior mean for any stock is the weighted average of its observed allele relative frequencies and the baseline center with the weights equal to simple ratios involving the baseline sample sizes and the constant of proportionality. The method is called pseudo-Bayes because it hedges by using the baseline samples both to choose the prior parameters and to evaluate the posterior distribution for the allele relative frequencies.

The Bayesian approach to describing the knowledge and uncertainty in the unknowns for complex problems is to generate a very large number of samples from their posterior distribution. Then summary measures of the posterior distribution, such as for location (mean, median, mode) and variation (standard deviation and various quartiles) can be computed from the samples with ignorable sampling error. In particular, the Bayesian method for stock mixtures used by Pella and Masuda (2001) is called a Markov chain Monte Carlo (MCMC) method because the samples for the unknowns are generated sequentially with each depending on the immediately preceding sample. The Bayesian method for stock mixtures also uses the data augmentation algorithm in which additional random observations, namely the unknown and purported sources of mixture individuals, are generated to greatly simplify estimation. At the *k-th* sample of the unknowns, let the current values from the posterior of **p** and **Q** be denoted by $\mathbf{p} = \mathbf{p}^{(k)}$ and $\mathbf{Q} = \mathbf{Q}^{(k)}$. The data augmentation algorithm has two steps:

1. Draw a random stock source of each mixture individual, $\mathbf{z}_m^{(k)} = \left(z_{m1}^{(k)}, \ldots, z_{mc}^{(k)}\right)$, where $z_{mi}^{(k)} = 1,\ i = 1, \ldots, c$, if the source of the *m-th* individual is the *i-th* stock, and $z_{mi}^{(k)} = 0$ otherwise. The stock source is drawn with *c* probabilities proportional to the posterior source probabilities based on the genotype and current values of **p** and **Q**.
2. Draw new values $\mathbf{p} = \mathbf{p}^{(k+1)}$ and $\mathbf{Q} = \mathbf{Q}^{(k+1)}$ from their respective posterior densities given the mixture sample genotypes, **X**, baseline samples for allele relative frequencies, **Y**, and the stock identities at step 1, $\mathbf{Z}^{(k)}$**.**

The posterior distribution for **p** is obtained by updating the Dirichlet prior for **p** with the assigned stock identities for the mixture individuals,

$\pi\left(\mathbf{p} \mid \mathbf{X}, \mathbf{Y}, \mathbf{Z}^{(k)}\right) = D\left(\mathbf{p} \mid \frac{1}{c} + \sum_{m=1}^{M} z_{m1}^{(k)}, \ldots, \frac{1}{c} + \sum_{m=1}^{M} z_{mc}^{(k)}\right)$. Notice that each updated Dirichlet parameter for a stock equals the sum of its prior parameter and the total number of mixture individuals assigned to the stock at the preceding sample in the chain. The posterior density for $\mathbf{Q}_i$ is $\pi\left(\mathbf{Q}_i \mid \mathbf{X}, \mathbf{Y}, \mathbf{Z}^{(k)}\right) = \prod_{h=1}^{H} D\left(\mathbf{p} \mid \beta_{h1} + y_{ih1} + \sum_{m=1}^{M} z_{mi}^{(k)} x_{mh1}, \ldots, \beta_{hJ_h} + y_{ihJ_h} + \sum_{m=1}^{M} z_{mi}^{(k)} x_{mhJ_h}\right)$. Notice that each updated Dirichlet parameter for an allele in the *i-th* stock equals the sum of its prior parameter, the count of the allele in the baseline sample, and the count of the allele for mixture individuals assigned to the *i-th* stock at the preceding sample in the chain.

The data augmentation algorithm generates a chain of samples from the posterior distribution for **p** and **Q**. However, the early samples in the sequence are influenced by the values chosen for **p** and **Q** to begin computations. Early burn-in samples need to be discarded, and a sufficiently large number of subsequent samples must be drawn to describe the posterior distribution accurately. Care is needed to assure that convergence has occurred. Recommended practice is to run several independent chains, ideally a chain for each reporting group, with dispersed starting points to reduce the possibility that a chain is accepted as representative of the posterior distribution before having converged. Usual starting points of the chains are mixture compositions **p** for which each reporting group is preponderant and each member stock contributes equally to the group. To determine the necessary chain lengths, an iterative scheme of processing pilot chains with the Raftery-Lewis convergence diagnostic (Raftery and Lewis 1996) is applied to each sequence of reporting group proportions (see Pella and Masuda 2001). To monitor convergence, the Gelman-Rubin shrink factor (Gelman and Rubin 1992) is computed for the mixture proportion from each reporting group. After convergence of chains is verified, the MCMC samples after burn-in are combined across chains. Various statistics of the pooled chains (equivalent to parameters of the posterior distribution given the large samples) such as means, standard deviations, and various percentiles (2.5, 5.0, and 97.5) are computed for the reporting group proportions. Also reported for each mixture individual are the chain average relative frequencies of assignment to each of the baseline stocks, which are the averages of the posterior distributions for posterior source probabilities and can be used to assign individuals to their sources by the MAP rule. Pella and Masuda (2001) provide the implementing software (see the Fortran program BAYES on the Auke Bay Laboratory website) for all computations. More recently, the algorithms have been reprogrammed at the Pacific Biological Station as the C program cBAYES, which is available at their website.

References for Appendix C

Alaska Department of Fish and Game 2003 SPAM Version 3.7: Statistics Program for Analyzing Mixtures. Anchorage, AK: Alaska Department of Fish and Game, Division of Commercial Fisheries, Gene Conservation Lab.

Debevec, E, R. Gates, M. Masuda, J. Pella, J. Reynolds, and L. Seeb 2000 SPAM (version 3.2): statistics program for analyzing mixtures. Journal of Heredity 91:509–510.

Fournier, D., T. Beacham, B. Riddell, and C. Busack. 1984. Estimating stock composition in mixed stock fisheries using morphometric, meristic, and electrophoretic characteristics. Canadian Journal of Fisheries and Aquatic Sciences 41:400–408.

Gelman, A., and D. B. Rubin. 1992. Inference from iterative simulation using multiple sequences. Statistical Science 7:457–511.

Kalinowski, S.T. 2003. Genetic Mixture Analysis 1.0. Bozeman, MT: Department of Ecology, Montana State University. Available at http://www.montana.edu/kalinowski.

Millar, R. B. 1987. Maximum likelihood estimation of mixed stock fishery composition. Canadian Journal of Fisheries and Aquatic Science 44:583–590.

Pella, J. and G. Milner. 1987. Use of genetic marks in stock composition analysis. *In* Population genetics and fisheries management (N. Ryman, F. Utter), pp. 247–276. Seattle, WA: University of Washington Press.

Pella, J. and M. Masuda, and S. Nelson. 1996. Search algorithms for computing stock composition of a mixture from traits of individuals by maximum likelihood. U.S. Department of Commerce, NOAA Tech. Memo, NMFS-AFSC-61, 68p.

Pella, J, and M. Masuda. 2001. Bayesian methods for analysis of stock mixtures from genetic characters. Fishery Bulletin. 99(1):151–167.

Raftery, A. E. and S. M. Lewis. 1996. Implementing MCMC. *In* Markov Chain Monte Carlo in practice: interdisciplinary statistics (W. R. Gilks, S. Richardson, D. J. Spiegelhalter editors). Chapman Hall/CRC.

Rannala, B. and J. L. Mountain. 1997. Detecting immigration by using multilocus genotypes. Proceeding of the National Academy of Sciences 94(17): 9197–9201.