Alaska Crab Economic Data Report Data Validation

2008 Calendar Year Data

Report Prepared for Pacific States Marine Fisheries Commission

November 2009

AKT LLP

Personal. Local. Global.

TABLE OF CONTENTS

Introduction 1

Methodology 3

Support Classes 5

Commendation 6

Appendix A 7

INTRODUCTION

BACKGROUND

The Bering Sea and Aleutian Islands (BSAI) Crab Rationalization Program was developed to create a quota system that grants exclusive harvesting and processing rights to crab harvesters, processors, and communities. The rationalized fishery began in fall 2005, with quota allocated to harvesters and processors based on historical participation in the fishery. Because of the expected impact on the industry, an economic data collection program was developed to better understand the economic impacts on the industry.

Economic data reports (EDRs) were developed to obtain information about the crab operations of harvesters and processors to help monitor how costs and economic returns of various stakeholders in BSAI crab fisheries are affected by rationalization. In order to ensure that the data submitted by respondents in the EDRs is accurate, Congress and the North Pacific Fishery Management Council specified that EDR data be subject to mandatory audits conducted by the third party collection agent, Pacific States Marine Fisheries Commission (PSMFC). PSMFC contracted AKT to develop and implement an EDR review and verification system, which involves reviewing the data contained within submitted EDRs, conducting verification audits for those EDRs containing data values outside of the expected range, and conducting random audits for a certain percentage of submitted EDRs.

The EDRs were developed to help determine the effects of the rationalization program, including changes to the costs of production and the effect of consolidation. National Marine Fisheries Service (NMFS) sought to understand the general trends over the years and the effects of rationalization to translate to other fisheries that are beginning similar programs.

This validation process is a continuation of similar work done in 2006 for EDR years 1998, 2001, 2004, and 2005, in 2007 for EDR year 2006, and in 2008 for EDR year 2007.

In summary, the purpose of the economic data report and data validation is to:

1) Aid the Council and NMFS in assessing the success of the Program.
2) Understand the economic performance of crab fisherman;
3) Understand how the economic performance has changed after rationalization;
4) Isolate the effects attributable to the crab rationalization program;
5) Assess the validity of data reported in submitted EDRs; and
6) Provide guidance on improvements in the EDR process to improve the validity of future data reporting.

KEY PARTICIPANTS AND ROLES

The key participants in the project include:

- *National Marine Fisheries Service (NMFS)* – initiator of the audit process and end-user of the information contained in the EDRs.

- *Pacific States Marine Fisheries Commission (PSMFC)* – collector and manager of the data collected through the EDRs.
- *AKT LLP* – independent accountants and consultants selected to audit and validate the information collected in the EDRs.
- Participants in the crab rationalization program.

Scope of Work

The following procedures were requested to be performed in the scope of work for this project:

1) ***Random Audits*** – Review and verification of a subset of data values reported in a randomly selected sample of EDRs.
2) ***Outlier Audits*** – Review and verification of data values reported in EDRs that contained multiple outlier variables. These outliers were identified through an analysis performed by NMFS.

The methodology to address the procedures above is outlined later in this report.

Based upon our conversations with NMFS and PSMFC, the key objectives of the audit were outlined as follows:

- Validate key data.
- Identify problems with the data or EDR instructions and make suggestions for future reporting.
- Promote compliance with timely and accurate data reporting requirements.
- Identify appropriate changes to data when missing or incorrect.
- Characterize, and in some cases quantify, the level of accuracy associated with particular data elements.

Key Information

The current analysis is based on the data collected from participants of the BSAI crab rationalization program for the year 2008. A statistical sample was determined based upon a total submitted population of 108, which was comprised of all unique submitters of information. The sample was determined based upon achieving a 95% confidence level with a precision level of 15% in terms of assessing the accuracy of the submitted data. (See Appendix A for detailed discussion of the statistical basis of the sample). The following table summarizes the number of EDRs submitted by type and the resulting sample size.

	# EDRs submitted	***Sample***
	2008	***2008***
Catcher Vessel	89	25
Processor (Catcher, Stationary Floating, Shoreside)	19	5

METHODOLOGY

AKT, PSMFC, and NMFS worked together to determine the best process to analyze data submitted through the EDR process and to determine the methodology to sample and audit the data submitted in the EDRs. The process was based on prior year experience with improvements made to benefit the participants. The following is a summary of the steps taken throughout the audit process.

1) ***Determine appropriate variables to validate.*** The significance of the data for random audits and available audit evidence were considered when determining the appropriate variables to validate. This is a collaborative process between PSMFC, NMFS, and AKT.

2) ***Determine population subject to random audit.*** The sample size was determined using a statistical model with a 95% confidence level and a 15% precision level. See Appendix A for a discussion of the statistical basis used for selection.

3) ***Determine outlier audit population and request information subject to audit.*** Based upon its analysis of the EDR data without vessel identity, NMFS identified the population that it desired to validate through an outlier audit. These audits focused on EDRs that had a significant number of outliers identified through analytical review. Once a vessel had been identified has an outlier, it was subject to validation for only those variables for which an outlier status had been identified. One vessel was selected as having outlier variables for the 2008 EDR data year.

4) ***Determining for-cause audits.*** Vessels selected for for-cause audits are those that did not comply with an audit request in the previous year. There were no for-cause audits for the 2008 EDR data year.

5) ***Gather and crosscheck EDR data to be audited.*** EDR data pertaining to the variables selected for auditing are transferred to AKT from PSMFC. AKT uses a standard auditing analysis spreadsheet and imports data from PSMFC into this format. AKT then verifies EDR data with the selected vessels original EDR submissions for accuracy.

6) ***Request information subject to audit.*** Selected vessels and processors are asked to provide supporting information for the variables selected for validation. They are given one month to comply with the request. Extensions are granted on an as-needed basis. If the selected vessels and processors do not comply within one month, they are individually contacted. Additional contact efforts are made as needed to ensure that each selected vessel and processor has an opportunity to respond in a timely manner.

7) ***Validate information by comparing with supporting documentation***. This process involves a review of the supporting documentation submitted against the original EDR data submission for each vessel and variable selected. Detailed notes related to the basis of information and quality of information is maintained in order to evaluate the validity of selected data. If clarifications on a discrepancy or additional supporting documents are needed, the vessel or processor is contacted as needed.

8) ***Summarize the results of the audit verification process***. Support categories were created to classify and summarize the validity of the audit evidence received. Each audited variable is classified within one of the support categories. This enables AKT to perform an overall analysis of the validity of the data.

AUDIT METHODOLOGY

AKT selects vessels or processors for random audit based upon the statistical sample outlined in Appendix A. AKT works with NMFS and PSMFC to determine the appropriate variables to validate.

For each data variable requested, AKT critically evaluates the support provided by the selected vessel or processor. Information is evaluated against third party support, such as invoices or fish tickets; internally-generated information, such as crew settlement sheets, general ledger details, invoices, detailed internal reports, or financial statements; and estimates made, including an assessment of the reasonableness of assumptions. Supporting documentation for internally-generated spreadsheets is requested on a judgmental basis. AKT also notes when no support is available to evaluate the information**.**

Many of the records provided to AKT are unique, especially for the vessels. The processor reporting is more formal and standardized, reflecting the large company nature of those operations. Because the material provided is so unique, the audit process begins with a detailed review of each information packet received while comparing totals for each variable to the original EDR entry. Each supporting document is assessed for accuracy and depth of support. Estimates are accepted as long as a reasonable explanation and/or calculation are also provided. Handwritten statements are also considered adequate, but only after discussion with the EDR preparer and requests for additional support.

AKT places phone calls to all submitters with estimates and hand written statements. AKT also validates all variables that are reported with no value (blank) or a zero value. If discrepancies are found between the original EDR submission and the supporting documentation, AKT contacts the vessel owner and/or preparer to validate the correct reported value. Many times this leads to receiving further documentation from the vessel and/or further explanation as to the methodology used to report EDR values.

If the initially provided documentation is determined to be insufficient support, or if support is missing for a certain variable, AKT contacts vessels to ask for further documentation. Once documentation is received, it is assessed and validated. In the current year, two vessels and two processors did not require follow-up information requests. All others required the follow-up research described above.

SUPPORT CLASSES

AKT worked jointly with PSMFC and NMFS to develop the following classifications to describe audit evaluations and summarize the results of the audited values.

Validation Code - Original Value	Original value substantiated	Audited value substantiated	Nature of Reporting Error	Correction	Validation Code - Audit Value
1	Yes	Yes (same)	No error; reported value is clearly substantiated by complete records	No	**1**
2	Yes	Yes (same)	Calculation error	Yes	**1**
3	Yes	Yes (same)	Misinterpretation of question	Yes	**1**
4	No	No	Reported value is "best guess"; value is not derived from records	No	**4**
5	Yes	Yes (same)	Data cannot be reported precisely as specified in EDR form and must be estimated; estimate is based on appropriate documentation and sound assumptions/logic and is considered validated	No	**5**
6	Yes	Yes (same)	Estimate is based on original documentation but flawed assumption/logic	Yes	**4**
7	Yes	Yes (updated)	Original value was reported correctly based on original documentation, but corrected based on updated documentation	Yes	**1**
8	No	Yes (new)	Original value is unsubstantiated; correction based on new documentation	Yes	**1**
9	No	No	No data reported/not applicable to vessel	No	**9**
9	Yes	Yes	Zero value is validated to be correct	No	**1**
9	No	Yes	Evidence contradicts a reported zero value.	Yes	**1**

The two basic populations that were sampled and evaluated based on the above criteria are Catcher Vessels and Processors (catcher, stationary floating, and shoreside). The records of 25 Catcher vessels were requested, and 25 were received. Six processors were requested to submit documentation, and five packets were received. All Catcher Vessels and Processors complied with AKT's requests for additional support, when applicable. Due to this high response rate, the support percentage is almost 100%. Accuracy of the original EDR data, however, varies greatly by vessel and by variable. This is especially true when one or two errors of larger size skew the result for the entire group.

One vessel was selected for outlier audits through the NMFS analysis process described earlier in this report. AKT requested support from this outlier for the unique variables that NMFS had selected for validation. This outlier complied with AKT's request for documentation, and AKT's subsequent request for additional documentation when applicable. Sixty-four variables were validated and supported for the selected outlier. Thirteen of the 64 original EDR values were corrected.

COMMENDATION

AKT worked collaboratively with members of the PSMFC and NMFS staff and would like to thank them for their commitment and time.

Name	*Organization*
Dave Colpo	Pacific States Marine Fisheries Commission
Geana Tyler	Pacific States Marine Fisheries Commission
Ron Felthoven	National Marine Fisheries Service
Brian Garber-Yonts	National Marine Fisheries Service
Audit participants	Individual vessels and/or processors

APPENDIX A

STATISTICAL SAMPLE

In order to determine an appropriate sample size as the basis of selection for the random audits, the main criteria to consider are the level of precision desired, the level of confidence or risk, and the degree of variability in the attributes being measured. These elements are defined as follows:

- ***Level of Precision*** – Also referred to as the margin of error, this is the range in which the true point value of the population is estimated to be. This is expressed as a percentage ± the true value (e.g., ± 5%). Thus, if it is found from the sample that on average 15% of the fisherman did not submit data then is could be concluded, that for the total population, between 10% and 20% of the fisherman have not submitted data.

- ***Confidence Level*** – The degree to which we are certain that a result, or estimate, obtained from the study includes the true population percentage, when the precision is taken into account. In a normal distribution 95% of the sample values are within two standard deviations of the true population value. If 100 vessels were sampled 95 would have the true population values within the range specified.

- ***Degree of Variability*** – This measures the variability within the population (e.g. Catcher Vessels, Catcher / Processor Vessels, Shore / Floating Processors, Large Vessels, Small Vessels). The more heterogeneous a population, the larger the sample size required to obtain a given level of precision. The more homogenous a population the smaller the sample size required. A variability of 50% signifies the greatest variability.

Due to the variability within the industry and the variability of the data being analyzed, there is not one specific variable that can be used to create a statistical model that would enable AKT to calculate a standard deviation and regression analysis for the project. This fact places the project in a similar category as a questionnaire, political poll, surveys, and extension program impacts.

While there are no statistical analyses that can be applied directly, there are similar projects that derive statistical sampling methods relating to extension program impact. In these projects the samples are used to evaluate a change made to the extension programs.

The sampling formulas derived for such projects and to ensure a statistical basis for the samples chosen are the following:

$$n_0 = \frac{Z^2(p)(q)}{(e)^2} \qquad n = \frac{n_0}{1 + \frac{(n_0 - 1)}{N}}$$

n_0 = Sample size
n = Sample size with finite population correction for proportions.
Z = The number of standard deviations a point x is from the mean. It is a scaled value.
p = Population variability.
q = 1 – p
e = The desired level of precision.
N = Total population.

For this project p (variability) equals .5 to account for maximum variability in the population.

This type of sampling methodology takes into account errors and missing information in the data. The precision level quantifies the tolerable level of error based on the sample size. This error level is then projected to the total population.

The samples were stratified based on the proportion of the group versus the total population. The reasoning behind this is that by sampling each individual population there would be no statistical basis for both the Catcher/Processor and Stationary/Floater Processors. The only way to have a statistical basis for this population would be to census the population. Because this is not a reasonable approach, AKT suggested that the population include all groups and then additional random audits be performed for the Catcher/Processor and Stationary/Floater Processor populations.

The sample population was ultimately chosen based upon a 95% confidence level with 15% precision and variability of 50% (due to the variability of the information requested). This method would ensure the data are correct (outlier audits) and it would also give a good idea for future projects how good the data is (random audits). This sampling method provides a statistical basis for future studies and gives the agencies a basis to measure the accuracy of the population data.