Report from Hawaii Bottomfish Commercial Fishery Data Workshops, 2015-2016

Annie Yau

U.S. DEPARTMENT OF COMMERCE
National Oceanic and Atmospheric Administration
National Marine Fisheries Service
Pacific Islands Fisheries Science Center

NOAA Technical Memorandum NMFS-PIFSC-68
https://doi.org/10.7289/V5/TM-PIFSC-68

February 2018

Report from Hawaii Bottomfish Commercial Fishery Data Workshops, 2015-2016

Annie Yau

Pacific Islands Fisheries Science Center
National Marine Fisheries Service
1845 Wasp Boulevard
Honolulu, HI 96818

NOAA Technical Memorandum NMFS-PIFSC-68

February 2018

U.S. Department of Commerce
Wilbur L. Ross, Jr., Secretary

National Oceanic and Atmospheric Administration
RDML Tim Gallaudet, Ph.D., USN Ret., Acting NOAA Administrator

National Marine Fisheries Service
Chris Oliver, Assistant Administrator for Fisheries

Recommended citation:

Yau, Annie. 2018. Report from Hawaii bottomfish commercial fishery data workshops, 2015–2016. NOAA Tech. Memo. NMFS-PIFSC-68, 105 p.

Copies of this report are available from:

Science Operations Division
Pacific Islands Fisheries Science Center
National Marine Fisheries Service
National Oceanic and Atmospheric Administration
1845 Wasp Boulevard, Building #176
Honolulu, Hawaii 96818

Or online at:

https://www.pifsc.noaa.gov/library/

Cover: Photo courtesy of Pacific Islands Fisheries Group.

Table of Contents

List of Tables

List of Figures

A. ACKNOWLEDGEMENTS

These workshops only succeeded because of the active engagement, hard work, dedication, time, and expertise of all the workshop participants. Special acknowledgement is given to fisher stakeholders Nathan Abe, Eddie Ebisui III, Roy Morioka, Layne Nakagawa, Basil Oshiro, Ed Watamura, and Leonard Yamada, who remained committed and active participants and who always participated on personal time and sometimes at personal expense to travel from different islands. Their valuable insights and contributions greatly improved the data resolutions. Many of the workshop participants spent countless hours before, during, and after the workshops to collate and analyze data and present results to the group. These included in alphabetical order Stefanie Dukes, Eric Fletcher, Dios Gonzales, Kimberlee Harding, Donald Kobayashi, Reginald Kokubun, Brian Langseth, Kimberly Lowe, Marlowe Sabater, Paul Tao, and Annie Yau with a heavier lift completed by Brian Langseth. Special thanks to Reginald Kokubun and NOAA staff Kimberlee Harding, who helped answer numerous questions about the history of DAR data reporting. Beth Lumsden and Chris Boggs did an excellent job recording detailed notes which were helpful in constructing this report. We also thank the many people who contributed with logistics to help the workshops succeed, including a special thanks to staff at Pier 38 for hosting 4 out of 5 workshops, and Council staff for providing space for the 4th workshop. Even if not all participants are mentioned by name in this paragraph (full list provided in Appendix C), the contribution of each person was vital to the collaborative and fruitful discussions.

B. INTRODUCTION

The challenges surrounding the robustness and utility for scientific and management purposes of the Hawaii Department of Aquatic Resources (DAR) commercial data have been known for some time. There are issues with both data quality and data structure that affect its utility for management and stock assessments. Reporting requirements and data forms have changed over time.

Scientists in the region have made multiple attempts over decades to refine DAR commercial data for stock assessment purposes. Starting at least in 1982, scientists have attempted to conduct stock assessments on Deep 7 bottomfish and, in doing so have applied analyses to address data quality and structure issues (Ralston & Polovina 1982). Many other attempts to resolve data issues have taken place over the iterations of stock assessments since then, and not all are described here. More recently, in 2004, the review, synthesis, and analysis of existing data were identified as a high priority at a series of stock assessment workshops convened to assess current bottomfish and coral reef fish assessment and data collection methods. Since then, various meetings continued to be held to discuss, identify, and resolve ongoing data issues, including a commercial catch-per-unit-effort (CPUE) workshop in 2008 (Moffitt et al., 2011). Overall, DAR has made efforts at improvements in data collection and the National Oceanic and Atmospheric Administration (NOAA) Pacific Islands Fisheries Science Center (PIFSC) has continued analyzing and exploring the DAR commercial data through multiple iterations of bottomfish stock assessments and related analyses.

Despite these ongoing efforts, this issue emerged again in December 2014, when NOAA contracted the Center for Independent Experts to conduct an independent external peer review of the 2014 main Hawaiian Islands Deep 7 bottomfish stock assessment which was completed by PIFSC scientists. This independent external review panel had strong reservations regarding the quality of the Hawaii DAR commercial Fisher Reporting System (FRS) data used to calculate input catch and CPUE for the stock assessment. The review panel concluded that although the methods to standardize CPUE and the assessment model were appropriate, the deficiencies in input data quality caused the stock assessment to have serious flaws that compromised its utility for management.

In response to these concerns and others, PIFSC organized a series of workshops bringing together participants from the NOAA Pacific Islands Regional Office (PIRO), the Western Pacific Regional Fishery Management Council (the Council), and fisher stakeholders to help understand and resolve these data issues conclusively. This was accomplished by identifying, and resolving, to the extent possible, issues in the DAR commercial data that affect the quality of data on Deep 7 bottomfish for stock assessment purposes. A series of workshops and small-group working sessions were held to facilitate completion of these tasks, with accompanying intersessional analyses. The ultimate goal was to produce a higher quality commercial data set that is deemed acceptable for stock assessment and management use by scientific, management, and user group (fisher) stakeholders.

Participants:

NOAA PIFSC and PIRO, the Hawaii DAR, the Council, and commercial fishers/stakeholders.

Scope:

Hawaii DAR data reporting on Deep 7 bottomfish or bottomfishing gear in the main Hawaiian Islands (MHI):

- Commercial Fisher Reporting System (FRS) Data, 1948–present
  This is the primary data set investigated, data through 2015 were used in workshop analyses.
- Dealer Reporting System (DRS) Data, 2000–present
  This secondary data set was investigated mainly to address issues pertaining to the FRS data set, data through 2015 were used in workshop analyses.

Goals:

(1) Identify and agree upon issues that affect the quality and consistency of DAR data on MHI Deep 7 bottomfish fishing.
(2) Investigate and agree upon an acceptable resolution to each issue, understanding that some issues may not be fully resolvable and/or may not have a significant effect on stock assessment results.
(3) Finalize a document describing the issues discussed and the agreed-upon resolution for each.
(4) Apply the list of resolutions to a copy of the DAR's commercial FRS data to create a stock assessment-suitable higher quality set of DAR FRS data on MHI Deep 7 bottomfish fishing for use by PIFSC in future stock assessments. This higher quality data set will constitute the best available scientific information for calculating total catch, weight, and CPUE of Deep 7 bottomfish species.

A total of 5 workshops spanning 15 days were held September 2015–November 2016, at conference rooms of NOAA Fisheries Pier 38 and the Western Pacific Regional Fishery Management Council office.

Workshop 1: September 2, 2015

Workshop 2: November 4–5, 2015

Workshop 3: April 26–27, 2016

Workshop 4: August 1–5, 2016

Workshop 5: November 14–18, 2016

All participants who participated in at least one workshop agreed upon the Terms of Reference (Appendix A). An acronym table and species table are provided in Appendix B. Workshop agendas and participant lists for each workshop are provided in Appendix C. **This report fulfills the third goal of finalizing a document describing the issues discussed and the agreed-upon resolution for each**, and this document was circulated to all workshop participants with comments incorporated. The report provides a comprehensive list of data issues, and captures the discussion and agreed-upon resolution to each issue. Generally, first a list of issues and possible proposed solutions was finalized at the workshops, and then some workshop participants conducted analyses between workshops to inform selection of resolutions when workshops were in session.

Participants have agreed that upon applying these resolutions to the Hawaii DAR commercial data, such a data set will constitute the best available commercial data for stock assessment purposes. Every attempt was made to seek and consider all currently available information in finding resolutions to each issue. While all of the issues in this report are known to exist in the DAR commercial data, not all of them have a major effect on stock assessment results and not all of them will be resolvable. Given the workload, available staff time, and the time frame within which outcomes were needed, participants retained the right to agree not to pursue finding a way to resolve some of these issues, especially when resolution was unlikely and/or the issue may have little impact on stock assessment results. Participants strived to investigate and agree upon an acceptable resolution to each issue, even if that resolution was to acknowledge that nothing can or should be done. Each issue listed here will not be revisited for future stock assessments unless additional information on this past data becomes available in the future.

Any resolutions that will continue to change DAR FRS data are documented here, so that if necessary, these can be consistently applied in the future. A list of errors in DAR FRS data will be provided to DAR. Additionally, this report can prompt discussion with DAR to evaluate whether long-term changes in the data collection forms and processes are warranted so that DAR can determine when and if such changes can be made at the source. If made, such changes would affect future collected data and so do not affect analysis of past data as outlined here. The data issues are organized according to three commercial data categories of interest, and then step-by-step from an assessment analysis perspective:

- Catch
- Fish weight (size)
- CPUE

An illustrative example of the DAR Commercial Fisher Reporting System (FRS) data structure using fake data is provided in Table 1 to help illustrate data structure and issues. The table does not provide a comprehensive list of all reported data fields, and only the most commonly used fields for stock assessment purposes are shown here. To emphasize, the data in Table 1 are fake, not based on real data, and represent fake entries for license, area, gear, port, and the use species name rather than code for illustrative purposes. Each row in the FRS data is a ‘record,’ which is a term used throughout this document.

Table 1. Illustrative example of some DAR Fisher Reporting System fields, using fake data to illustrate data structure and data issues.

record	LICENSE	DATE	AREA	GEAR	HOURS	SPECIES	#CAUGHT	LBS	PORT
1	649XX	9/15/1984	105	3	0	Onaga	10	31	0
2	649XX	9/15/1984	105	3	0	Gindai	2	17	0
3	649XX	9/15/1984	105	3	0	Hapu	1	200	0
4	649XX	11/2/2015	344	3	8	Opaka	12	35	301
5	649XX	11/2/2015	402	99	8	Mahi	3	6	400

Some of the more commonly reported fields are found in Table 1 including commercial marine license number, date, area code, gear code, hours fished, species, number caught, pounds caught, and port code. There are illustrations of several data issues in this table, including data quality, lack of an identifier for a unique trip, lack of targeting information, and lack of information on individual fish sizes. These are just some of the data issues tackled by the participants.

Data quality is a question for records 1–3, in which a fisher reported 3 species of bottomfish caught in one area on one date but reports a 200-lb Hapuupuu, which is about twice its maximum known size. It is unclear if this is a data reporting or recording error, but it is unlikely a valid entry. Records 4–5 illustrate the challenges of lacking a unique trip identifier or targeting information. A fisher reported catching 12 Opakapaka in one area and then 3 Mahimahi in an area far away using a different gear on the same day. It is impossible to know whether the fisher caught both fish in the same trip, or whether the fisher returned to port in between. The field "HOURS" fished was only available after a reporting form change in 2002, but the number of hours fished in this instance does not provide enough additional detail to help discern this detail. Also it is impossible to know if this fisher was targeting Opakapaka or Mahimahi, or both. Given the distance between the grids reported, it is also possible that these fishing records occurred on two different days but were reported on the same day. Finally, the records have data on the total number caught and the total pounds of that number. Individual fish sizes (weights) are only available for records that report 1 fish of 1 species, otherwise only an average fish size is available per record.

The primary relevant field for calculating commercial catch is "LBS." For CPUE, the primary relevant field for calculating commercial effort is "HOURS" fished, although a different effort metric needs to be applied prior to October 2002 because that field was not reported historically. For fish weight (size), the primary relevant fields are "#CAUGHT" and "LBS." Other fields are also used to explore data issues and apply filters.

In the process of conducting these workshops, the participants addressed several priority research topics developed by a 2015 workshop on Bottomfish Research (Table 1 in Yau & Oram, 2016). The priorities addressed were: priority 5, to explore further variables in CPUE standardization; priority 10, to identify data collection and education needs; and priority 17, investigation of commercial fish size data.

C. SUMMARY OF WORKSHOP OUTCOMES

C.1. Summary of Catch Resolutions

The purpose of catch is to provide the total annual commercial fishery removals/deaths from the stock.

<u>Final method to calculate total commercial catch:</u>

Add up the number of pounds of each Deep 7 species reported as caught, regardless of which gear was used to capture them.

To implement using the DAR FRS data:

(i) Select all records that are for areas in the main Hawaiian Islands (Figure 1, Table 2).
(ii) Select all records where the species code indicates a Deep 7 bottomfish (Table 12).
(iii) For each record, take the pounds as the maximum of "POUNDS_KEPT" and "POUNDS_SOLD" and sum up this value for all records in a year to calculate annual total commercial catch. This mainly applies to records prior to October 2002, after which "POUNDS_SOLD" is no longer consistently reported by fishers but by dealers instead. Summing up according to step (iii) may also be done for individual species to calculate species annual total commercial catches.

C.2. Summary of Fish Weight (Size) Resolutions

<u>Fish weight (size) is used to</u> determine changes in fish size indicated by annual average commercially caught individual fish weight.

<u>Final method to calculate annual commercially caught fish weight:</u>

(i) Use the Fisher-Reporting System (FRS) data, not the Dealer-Reporting System (DRS) data.
(ii) Select all records that report a Deep 7 bottomfish species (Table 12).
(iii) Select all records that are for areas in the main Hawaiian Islands (Figure 1, Table 2).
(iv) Exclude records up to June 30, 1948, because there is not data available for that full fiscal year 1948.
(v) Select only records that reported using bottomfishing gear (gear type = 3, deep sea handline).
(vi) Exclude records reporting either 0 pounds, or 0 numbers of fish.
(vii) Exclude records where the reported average fish weight is either greater than a species-specific maximum or less than a species-specific minimum (Table 4).
(viii) For an individual species, use all records to calculate annual average fish weights and associated error.
(ix) For a Deep 7 complex-level fish weight time series, use an annual average fish weight and associated error that combines all Opakapaka and Onaga records.

C.3. Summary of Catch-per-unit-effort Resolutions

A catch-per-unit-effort that accurately reflects the trends in CPUE over time can be standardized and used as a relative abundance index in a stock assessment.

Final method to calculate commercial nominal catch-per-unit-effort:

CPUE Step 1: Selecting data records considered as bottomfish records

(i) Select all records that are for areas in the main Hawaiian Islands (Figure 1, Table 2).
(ii) Select only records that reported using bottomfishing gear (gear type = 3, deep sea handline).
(iii) Retain duplicate records.
(iv) Remove records with license = 0.

Getting at records that target bottomfish:

(v) For trips prior to October 1, 2002, remove 'single reporting days' that report gear type 3, catch 0 pounds of Deep 7 bottomfish, and catch any of the following:
- Any pelagic species (Table 12**).**
- Uku (species code = 20)
- Species code = 0 (an invalid code)

(vi) Remove records that occurred on dates between when the Deep 7 bottomfish fishery was closed because the catch limit was reached and the end of the fishing season. Closures occurred from April 16, 2008 to August 31, 2008; July 6, 2009 to August 31, 2009; April 20, 2010 to August 31, 2010; and March 12, 2011 to August 31, 2011.
(vii) Remove 'single reporting days' that caught less than 50 lb of Deep 7 AND caught any pelagic species (Table 12) from 1948 to 1985 for areas around Kona and South Point (see Figure 18, grid area codes 102, 122, 101, 121, 100, 120, 108, 128) since it's likely they were palu ahi fishing rather than Deep 7 bottomfishing.

CPUE Step 2: Select data records that comply with a unit of effort

From 1948 to Sep 2002:

(viii) Use 'single-reporting day' as an effort metric from 1948 to Sep 2002, defined as a unique combination of license number and date.

Account for multi-day trips showing up as a single day in this data by applying the following:

(ix) For 'single reporting days' that report one area and two ports, divide into 2 fishing days if the two ports are on nearby islands.
(x) For 'single reporting days' that report multiple areas and multiple ports, separate into multiple fishing days based on number of ports.
(xi) For all of the above 'single reporting days' that report both area and port, divide into 2 fishing days if distance from area to port is between 30.1 and 60 nm; divide into 3 fishing days if distance from area to port is between 60.1 and 90 nm; and so on.
(xii) For records prior to October 1, 2002, remove all records from licenses that only ever reported on the first and/or last day of the month.

From Oct 2002 to present:

(xiii) Use hours fished as an effort metric starting from Oct 2002.

CPUE Step 3: Select data that are representative of CPUE trends over time.

(xiv) Remove records from fishers who used gear 3 but never caught any pounds of Deep 7.
(xv) Remove single-reporting days that occurred during fishery-independent sampling for specific fishers in specific grids.

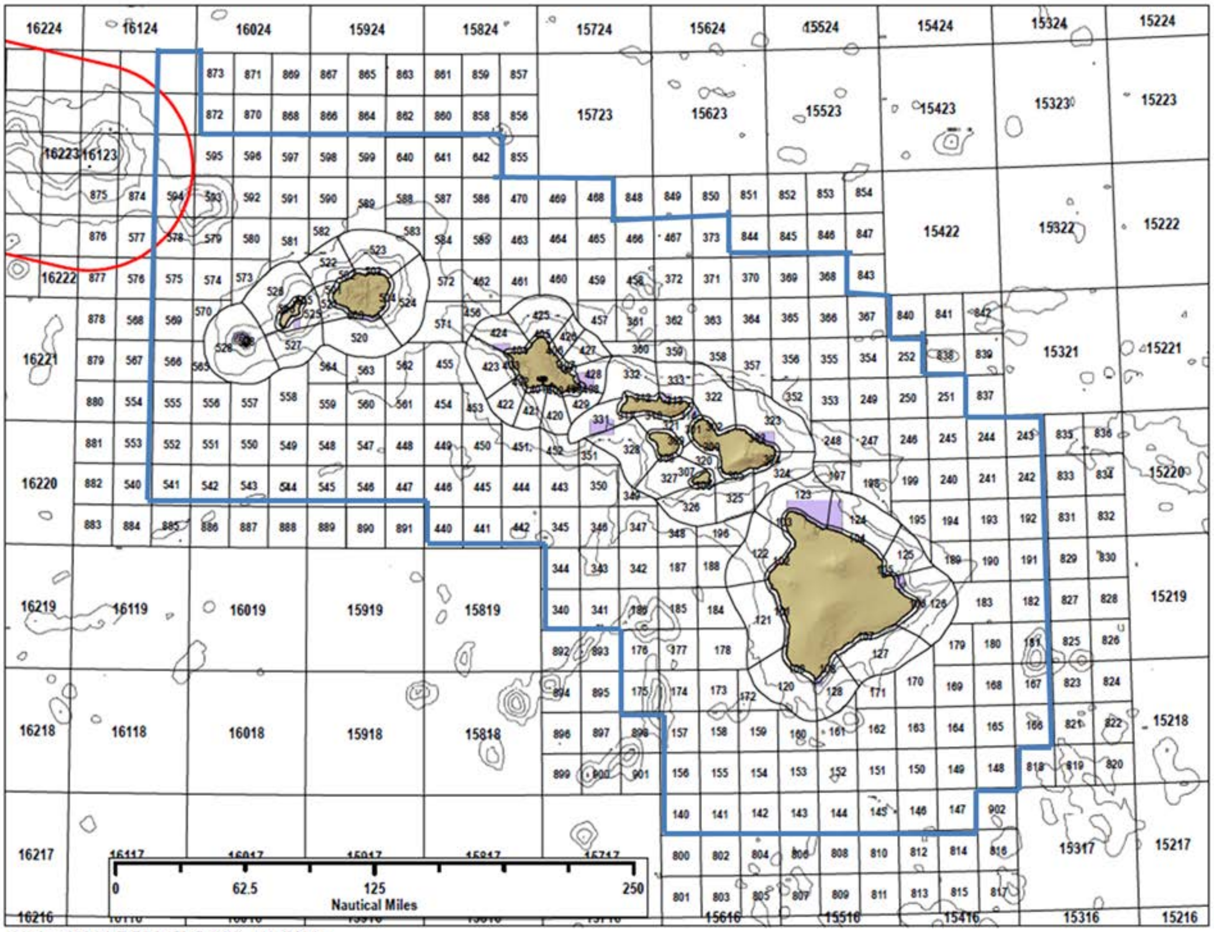
CPUE Step 4: Standardize CPUE to account for factors other than changes in abundance.

This step is part of stock assessment modeling and not data filtering, so was out of scope for the workshop. However, the workshop participants came up with a prioritized list standardization factors to explore as specified in detail below in Section D.3.4.

D. WORKSHOP DISCUSSIONS

What follows are detailed descriptions of the issues, discussions, analyses, and resolutions for the three major data categories. Some generic data issues such as missing fields, erroneous/out of range values or codes, and duplicate records are addressed under each data category. The document by Moffitt et al. (2011) provides an excellent summary of commercial DAR FRS and DRS data, descriptions of data fields available, and history of the bottomfish fishery and was referred to in detail during workshop analyses. Code that can be used to implement the agreed-upon resolutions for all 3 data categories is available in Appendices D.1–D.3.

For the upcoming benchmark stock assessment scheduled to be delivered in 2018, and thus for analyses presented here, records up to June 30, 2015 (state fiscal year 2015) were used. For each of the three data categories, the main Hawaiian Islands was defined as presented in Figure 1, with corresponding list of DAR area codes in Table 2. The exact area comes from a DAR definition as proposed by Reggie Kokubun during the workshop and modified by workshop participants to exclude grid areas 16024, 16025, and 16026 because no Deep 7 bottomfish were caught in these areas and these areas were not contiguous with the rest of the map. The final definition was agreed upon after Reggie Kokubun found that 323 out of a set of 605 area codes PIFSC had on file and had planned to use were invalid for bottomfish; some referred to fish ponds, and others did not correspond to a physical grid area. The remaining grid areas all correctly correspond to a MHI grid area. It was also found that the final definition differed from definitions used in previous assessments (219 records would be omitted, and 60 would be added if moving from the map in previous assessment to the new map). Given that differences in records included were small, and that the current definition aligned with the boundary of the NWHI Mau Zone (161°20’ W) as stated in the federal registry (54 FR 29907, September 6, 1988), the map was adopted.



Source: NOAA PIFSC IFMP (MKL 11/1/2016)

Figure 1. Boundary of the main Hawaiian Islands (blue outline), with the Papahānaumokuākea National Monument as of August 25, 2016, prior to subsequent expansion in red.

Table 2. Area codes by DAR that consist of the main Hawaiian Islands.

100	158	194	322	366	448	525	579
101	159	195	323	367	449	526	580
102	160	196	324	368	450	527	581
103	161	197	325	369	451	528	582
104	162	198	326	370	452	541	583
105	163	199	327	371	453	542	584
106	164	240	328	372	454	543	585
107	165	241	331	373	455	544	586
108	166	242	332	400	456	545	587
120	167	243	333	401	457	546	588
121	168	244	340	402	458	547	589
122	169	245	341	403	459	548	590
123	170	246	342	404	460	549	591
124	171	247	343	405	461	550	592
125	172	248	344	406	462	551	593
126	173	249	345	407	463	552	594
127	174	250	346	408	464	555	595
128	175	251	347	409	465	556	596
140	176	252	348	420	466	557	597
141	177	300	349	421	467	558	598
142	178	301	350	422	468	559	599
143	179	302	351	423	469	560	640
144	180	303	352	424	470	561	641
145	181	304	353	425	500	562	642
146	182	305	354	426	501	563	16123 I
147	183	306	355	427	502	564	16123 F
148	184	307	356	428	503	565	16123 C
149	185	308	357	429	504	566	
150	186	309	358	440	505	569	
151	187	310	359	441	506	570	
152	188	311	360	442	508	571	
153	189	312	361	443	520	572	
154	190	313	362	444	521	573	
155	191	314	363	445	522	574	
156	192	320	364	446	523	575	
157	193	321	365	447	524	578	

D.1. CATCH

Catch is used to provide the total annual commercial fishery removals/deaths from the stock.

Issues:

D.1.1. Discarded, Released, and Predated Catch

Although likely to be a rare occurrence, bottomfish may be discarded, released or predated. For any information on such interactions that result in mortality (such as for predated fish), it is best to include this information in estimates of annual commercial catch. Some data fields in the DAR FRS data starting in October 2002 provide some information about this issue: “LOST” (number of fish lost due to predators) and “RELEASE” (number of fish released dead or alive).

It is difficult to always tell when a fish has been predated. The presence of predators is one factor; feeling a tug and then slack line could be another indicator. Discards are rare. Fish reported as released are likely released as part of scientific tagging studies, and not as part of the commercial fishery. It is not clear the extent to which all fishers are reporting these interactions, how they choose to report, and how consistent is the reporting. Noting that it can be difficult to determine what species of fish was lost or released if it was not seen at the surface, species-level reporting of these fields is not required for stock assessment purposes, and Deep 7 bottomfish totals are sufficient. Most fishers will know whether they lost or released a Deep 7 bottomfish, even if they aren’t sure of the exact species. Fish that are predated are clearly removed from the fish stock, but it is unclear whether discarded and released fish die shortly thereafter given the potential for barotrauma due to the deep depths from which bottomfish are retrieved. Only fish that die need to be included in the calculation of annual commercial catch.

The information on this issue is limited, starting only in October 2002. Only a few data entries provide information in these fields. Additionally, for purposes of calculating catch, these records would ideally be provided in terms of weight such as pounds. However, the “LOST” and “RELEASE” fields are reported in numbers of fish. This may be in part because it is very difficult to guess the weight of fish that are lost or released.

From 2002 to 2015, anywhere from 10 to 20% of records reporting Deep 7 bottomfish catch also reported Deep 7 bottomfish “LOST” to predation. This percentage was on average slightly higher for Niihau and Oahu (~16%) compared to Big Island and the Maui Nui area. The Big Island had on average the lowest percentage of records reporting “LOST”; less than 10% of records.

From 2002 to 2015, anywhere from 1 to 12% of records reporting Deep 7 bottomfish catch also reported Deep 7 bottomfish “RELEASED.” On average, the highest percentage came from Maui Nui (~7%) and fishers noted this is likely a result of tagging efforts. Niihau (~6%), Oahu (~4%), and the Big Island (~3%) had smaller average percentages of Deep 7 bottomfish records reporting “RELEASED.”

The number of Deep 7 bottomfish reported lost or released varies by species and by year. From 2006 to 2014, on average the annual number recorded as lost is 2.4% of the number caught. From 2006 to 2014, on average the annual number released is 1.1% of the number caught.

RESOLUTION for Discarded, Released, and Predated Catch:

DO NOT include the numbers discarded, released or predated ("LOST" and "RELEASE") in calculation of total commercial catch of Deep 7 bottomfish. Compared to the total magnitude of commercial catch and estimated unreported catch, the numbers discarded, released, and predated are very minor.

D.1.2. Underreporting of Commercial Catch

Previously, for many years, fishers used dealer tickets to report only sold catch, and not necessarily catch given away, bartered, taken home, or otherwise not reported by a dealer. There was and continues to be underreporting of commercial catch. Recognizing that a) the DAR FRS data do not fully capture non-commercial catch such as subsistence catch, recreational catch, depredated catch, and catch used for barter, and b) the currently used stock assessment includes unreported catch separate from and in addition to DAR reported catch, workshop participants investigated and discussed the degree to which commercial fishers underreport commercial catch.

In the more distant past, it was true that fishers often used dealer tickets (sales records) to fill out fisher reports. One reason is convenience; another is that dealers come up with exact weight measurements whereas fishers are often estimating, although estimating fairly accurately. In recent years, fishers have started to consistently report more than just what was sold according to dealer ticket. Dealer data for bottomfish begins in 2000, so prior to that there is no direct way to compare dealer and fisher reports, and it will be difficult to determine how much underreporting occurred prior to 2000.

According to FRS data from 1948 to September 2002, which includes fields of pounds kept and pounds sold, fishers reported on average 3.4% more pounds sold than kept. The pounds kept and pounds sold values are reasonably close during this time period when it was known that fishers used dealer receipts to fill out pounds sold (and possibly also pounds kept as well). Since 2002, this ratio has increased. From 2007 to 2015, fishers report on average 19.8% more fish by weight than dealers report as sold. This suggests that in recent years, fishers have been reporting fish caught that do not go to registered dealers. This makes sense given that in recent years, managers and scientists have emphasized to the fishing community the importance of reporting of all catch, which may have contributed to more complete reporting. It's likely that prior to 2002, the DAR fisher-reported commercial data is underreporting commercial bottomfish catch, but there are no other data sources to help determine how much underreporting was occurring.

Workshop participants agreed that commercial fishers underreported bottomfish in the past especially prior to 2002, and that fishers caught at least 20% more than they reported, but likely more than this. One reason is that the price of fish is higher today than historically, so prior to 2002, fish were more often given away.

Starting in the 2011 stock assessment (Brodziak et al., 2011) and continuing in the 2014 stock assessment (Brodziak et al., 2014), estimates of unreported commercial and recreational Deep 7 bottomfish catch were included as summarized by Courtney and Brodziak (2011). These estimates were based on a series of studies, each of which attempted to quantify the ratio of unreported to reported catch at different points in time. The methods for each of these studies were checked in detail and underreporting of commercial catch is included in the estimates of unreported catch from these studies. Since unreported commercial catch, along with recreational catch is already being accounted for in another part of the assessment, workshop participants agreed not to try and account for underreporting of commercial catch in another way.

RESOLUTION for Underreporting of Commercial Catch:

The group recognized there is no rigorous quantitative way to account for underreported commercial catch given the current data structure. Given that the previous 2011 and 2014 stock assessments already accounted for underreporting of commercial and recreational catch through inclusion of an estimated ratio of unreported to reported catch estimated from several different studies, and as long as such an accounting of underreporting of commercial and recreational catch continues to be used, then there is no further need to adjust the commercial catch in another way to account for underreporting of commercial catch.

D.1.3. Misidentified Species Catch

There is occasional misidentification of one Deep 7 bottomfish species for another. This does not affect calculation of total Deep 7 annual commercial catch but may affect individual species annual commercial catch if such numbers are calculated and used in the future.

There is no way to determine whether a species was misidentified unless the dealer-reported and fisher-reported information differs. Even then, species misidentifications will be difficult to detect given that fish are not reported individually but grouped (ostensibly by species) by both dealers and fishers when reporting. Misidentification should be less of a problem in the last ~15 years. No participant expressed concern that there could be an appreciable frequency of misidentification of the principle target species.

RESOLUTION for Misidentified Species Catch:

Misidentified species catch is a minor issue and DOES NOT need to be investigated.

D.2. FISH WEIGHT (SIZE)

Fish weight (size) is used to determine changes in fish size indicated by annual average commercially-caught individual fish weight.

Fish weight data, to characterize fish size, has not been included in a Deep 7 bottomfish stock assessment used for management purposes in recent memory. PIFSC scientists spent time introducing the concept of using average and individual fish weights to indicate fish size in stock assessments, which can be used in addition to catch and CPUE but can also be used alone for example for calculation of life history parameters. Fish weights need to represent what is happening to the sizes of fish in the stock, in the water. Both the fish weight data trend (increasing vs. decreasing sizes) and scale (exact sizes in pounds) are important. Information on fish weights over time is used to infer changes in the stock over time including age, mortality, and recruitment. Thus, changes in fish weights over time as illustrated by fishery data should generally reflect changes in the actual weights of fish in the stock, over time. In other words, changes in commercial fish weights over time should *not* be driven largely by targeting, market forces, or any other factors (although the influence of these factors can be present, just not dominant). Participants were asked to select data on fish weights that is accurate, reliable, and representative of fish weights for stock assessment use.

All participants, especially the fishers, had numerous questions and discussion about how fish weights are used in stock assessments and what it means if fish size increases or decreases over time. One fisher said the size of individual fish can be driven by many forces and not just changes in the stock's average fish size. Questions were raised about the extent of life history knowledge on all Deep 7 species, all relating to maturity and the contribution of a fish to the stock based on its size. Most of the life history studies have focused on Opakapaka, with Hapuupuu as a secondary focus and Onaga the next priority. However there is some life history information available for all Deep 7 species. Participants struggled with what it means to have a combined complex-level weight time series, how to come up with a representative fish weight time series for a complex, and if it even made sense to calculate a complex-level time series given that the majority of participants were uncomfortable with the idea of combining life history parameter values.

Participants from all sectors voiced dissatisfaction at the current fisher-reporting system, which has fishers estimate the total weight for a species and then report the corresponding number of fish. This means only average weights can be calculated and individual fish weights are not always available, and many participants had questions about whether average fish weights could still be useful metrics for monitoring changes in the stock and fishery. A concern brought up by several fishers was whether the signal from a large fish and thus its importance would be swamped out in an average. This concern stems from the disproportionately higher contribution of such large fish to overall reproduction of the stock. PIFSC scientists spent time explaining how weighted averages are calculated, and provided exercises illustrating how a single large fish among many could change the resulting average fish weight.

Fishers report that size targeting occurs and their estimate was that it is successful about 80% of the time.

<u>Issues:</u>

D.2.1. Selecting Which of the Two Fish Weight Data Sets to Use

Both the DAR FRS and DRS data sets include information about total weight and number of fish caught, and both data sets were considered. Fisher-reported data on fish weights includes recognized biases that affect whether the weight data is representative of weights caught by the fishery. Many of these are targeting biases. Fishers are known to target specific sizes depending on market demands even if they do catch fish of all sizes throughout the year. Fishers report that bottomfish will school according to size so that it's possible to target size. Fishers generally target larger bottomfish. Around the holiday season in winter, fishers target smaller ones to be able to give away more pieces. Location and distance to port affects sizes caught as well. Also, not all fish caught are reported in the fisher-reported data. If only a few fish are caught, and they're small, it is possible fishers don't bother reporting.

Advantages of the FRS data set are that data is available from 1948 to present day, and the weight data also has species, gear type, and area code information. Disadvantages of the FRS data set are that fishers estimate total weight of one species caught and also report total number caught. This results in fisher-reported weights being estimates, and that only average weights can be calculated rather than individual weights. Information about individual fish, especially ones that are largest or smallest, is lost because only average weights can be calculated.

There are recognized biases for the dealer-reported weights as well. Again, Hawaii state regulations require bottomfish caught to be at least 1 lb so fish smaller than this will not show up at the dealer to be sold. The United Fishing Agency purchases all fish meeting this state requirement. Some other dealers, especially Chinese restaurants, prefer fish < 2 lb in weight as these will fit whole on a platter. These dealers will even pay higher prices for fish < 2 lb in weight. Open air markets sell fish whole, and won't buy fish > 5 lb. Gindai is not often sold at market, but is generally eaten by the fisher or gifted due to its smaller size and flavor. It's very possible that fish kept for home consumption and fish given away as gifts are not recorded in the dealer data. All of these practices result in biases in the DRS weight data because not all fish that are caught are actually sold.

The advantages to the DRS data are that the dealer measures rather than estimates the weights of fish. When several fish of the same species are grouped together by the dealer, they are often grouped according to size. Size variation within a dealer lot is low compared to variation between dealer lots (Ralston et al., 1986), although it is not clear how consistent this is over time since a study was only done in 1986. Calculating an average fish weight from dealer-reported data on total weight and numbers is thus more reliable than calculating an average fish weight from fisher-reported data because of the more accurate weight measurement and the grouping by size. The dealer tends to separate the largest fish into individual lots for sale, allowing for accurate weight reporting of the largest fish caught. Disadvantages of using the DRS data on weights are that the data only starts in 2000 (the United Fishing Agency may have records prior to this, but such records are not publicly available), and that the location of fish caught is not recorded in the dealer data so you cannot filter out bottomfish caught in the Northwest Hawaiian Islands (that fishery continued until 2009). Fish also lose water weight and shrink between the time they are caught and the time they are delivered to and measured by the dealer.

Participants considered both the FRS and DRS data sets for weight, and compared the two in terms of number of fish reported and the mean average weight (the weighted average of the average weight per record) (Figure 2). Both the number and mean average weight of Deep 7 bottomfish reported in the DRS data set declines after 2009 because fishing in the Northwest Hawaiian Islands is banned thereafter. The participants noted that the inability to select only main Hawaiian Islands data from the DRS is a serious disadvantage to that data set.

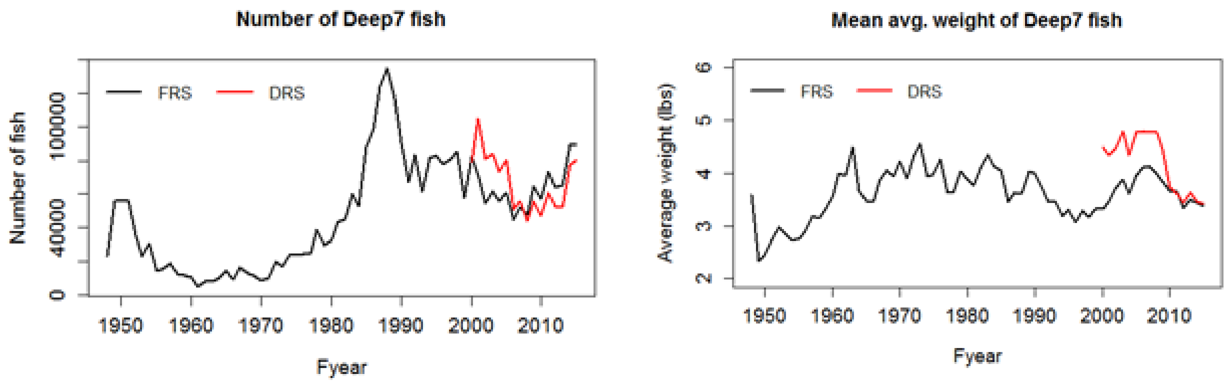


Figure 2. (Left) The number and (right) mean average annual weight of Deep 7 bottomfish reported in the FRS (black line) from 1948 to 2015 and in the DRS (red line) from 2000 to 2015, the latter which includes fish from the NWHI.

In the interest of being thorough, workshop participants continued to explore the DRS data set on weight. A noted disadvantage of both the DRS and FRS data sets is the grouping of multiple fish in a single record, allowing only an average weight to be calculated. Participants found this grouping to be unfortunate, since it masks individual signals of fishes at the extreme sizes (large and small). Small fish could indicate a recruitment pulse and large fish are disproportionately more important as egg producers, but neither's contribution would be fully recognized in an analysis if they were grouped with other fish. One suggestion was to investigate whether only 1-fish records should be used in the assessment, since they represent measured weights of individual fish rather than aggregated weight of several fish. Participants quickly discovered there are differences in the patterns from 1-fish vs. all fish records. Quite a few records are 1-fish records with Hapuupuu and Onaga being sold the majority of the time as 1-fish records (Figure 3). If only 1-fish records in the DRS data set are analyzed, then most of them are Onaga whereas Opakapaka is the dominant species when all records are analyzed (Figure 3). Overall, 47% of Deep 7 DRS records report only 1-fish, and 86% of records report between 1 and 5 fish (Table 3). Notably, the mean average annual weight for Deep 7 bottomfish as reported in the DRS data set varies depending on which records (1-fish, 1-2 fish, up to all fish records) are used (Figure 4). When only1-fish records are selected, the mean average annual weights are higher. Presumably this is an effect of larger fish being sold as individual fish by the dealer.

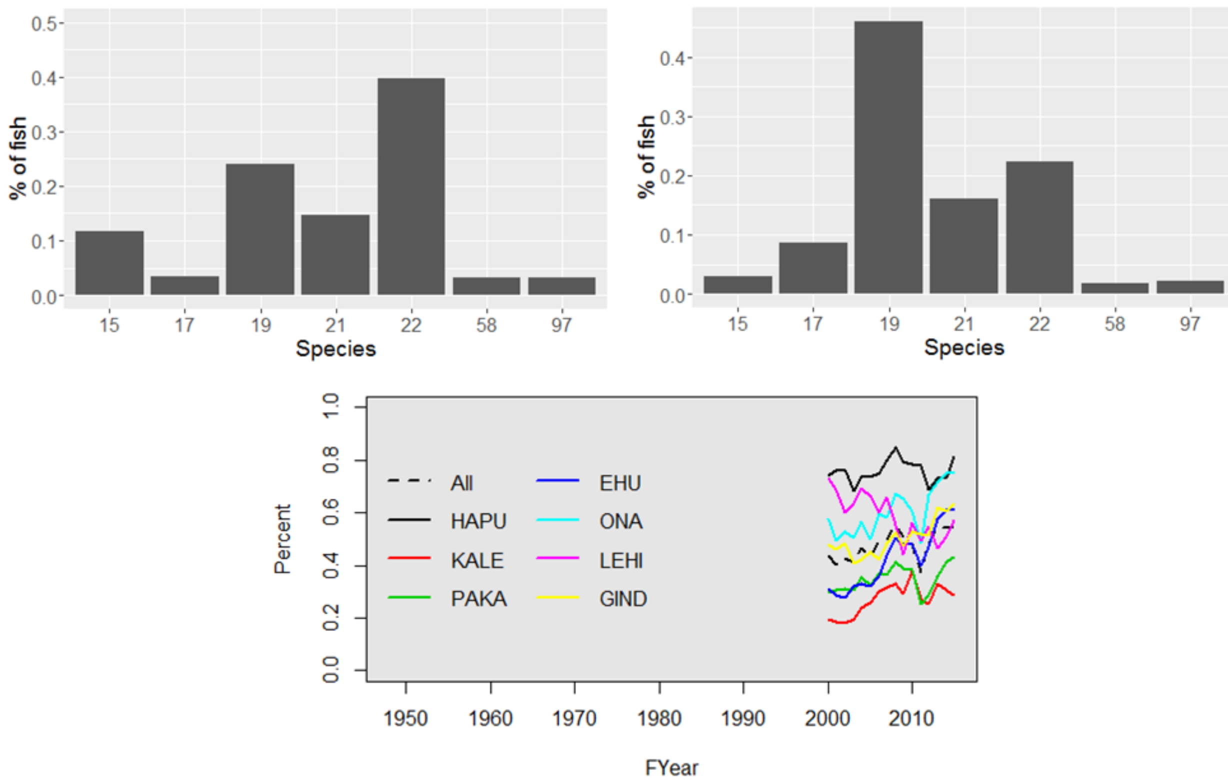


Figure 3. Species composition by weight of DRS records reporting weights of (top left) 1-fish, and (top right) all fish. (Bottom) The proportion of all DRS records for a given species that are reported as 1-fish records. Species are referred to by the first few letters of their common name or by species code.

Table 3. The percentage of Deep 7 bottomfish records reporting weight according to different numbers of fish reported in the DRS data.

Data Subset	Percentage of Records	Percentage of Fish Numbers Reported
1 fish	47	13
1–2 fish	62	21
1–3 fish	75	31
1–4 fish	80	36
1–5 fish	86	45
5+ fish	20	64
All fish	100	100

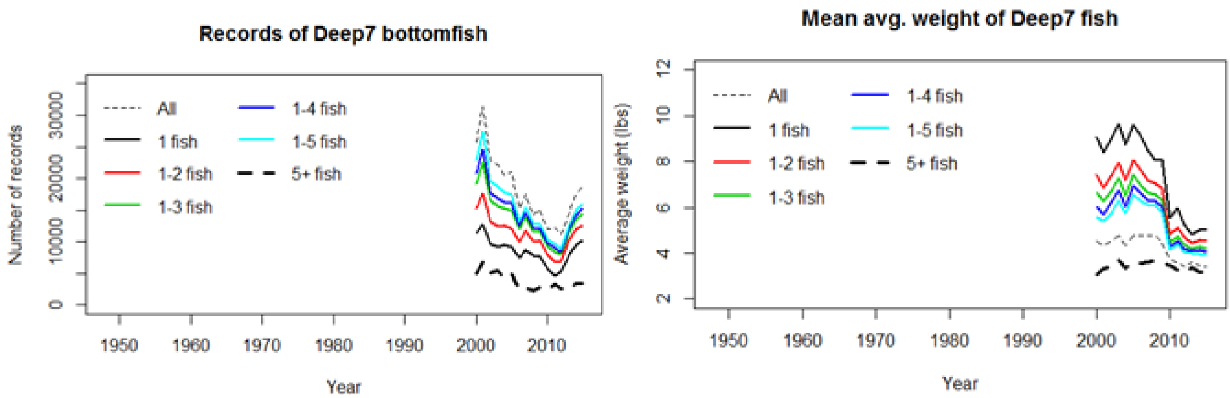


Figure 4. (Left) Number of records in DRS data set reporting any Deep 7 bottomfish species for various combinations of fish numbers per record. (Right) Average annual weight of Deep 7 bottomfish as reported in the FRS data set for various combinations of fish numbers per record.

The same pattern is evident when the Deep 7 are separated into individual species. The average weight of fish reported in 1-fish records is higher than the average weight of fish reported in all records for each species, as evidenced in plots of the mean annual weights by species (Figure 5). Number of records and mean average annual weight trends by species are plotted in Figure 6.

Participants decided NOT to use data from the Dealer-reporting System for calculating weight for stock assessment purposes. They recommended continuing to collect this data, and use it for other purposes including but not limited to investigating market trends and drivers and individual sizes of fish including details of larger fish caught. DRS data is still useful for comparing to FRS data to answer other research questions, such as comparing the average weight that goes to market vs. average weight caught. There was a suggestion to compare the 1-fish records in the DRS data to the 1–2 fish records in the FRS data, to see where the split from 1-fish to multiple fish occurs. The DRS data also includes more 1-fish records that can be used to calculate a size distribution (although it may be biased) than the FRS data because of the way dealers separate out fish for sale.

The decision to use FRS and not DRS data for weight calculation was driven by several factors: the dealer data cover a shorter more recent time period, much of the dealer information is already in the fisher-reported data, fish from the NWHI could not be filtered out, and not all commercially caught fish make it to dealers. This decision was unanimous.

RESOLUTION for selecting whether to use one or both of the two weight data sets available to characterize fish size:

(i) Use the Fisher-reporting System (FRS) data, not the Dealer-reporting System (DRS) data.

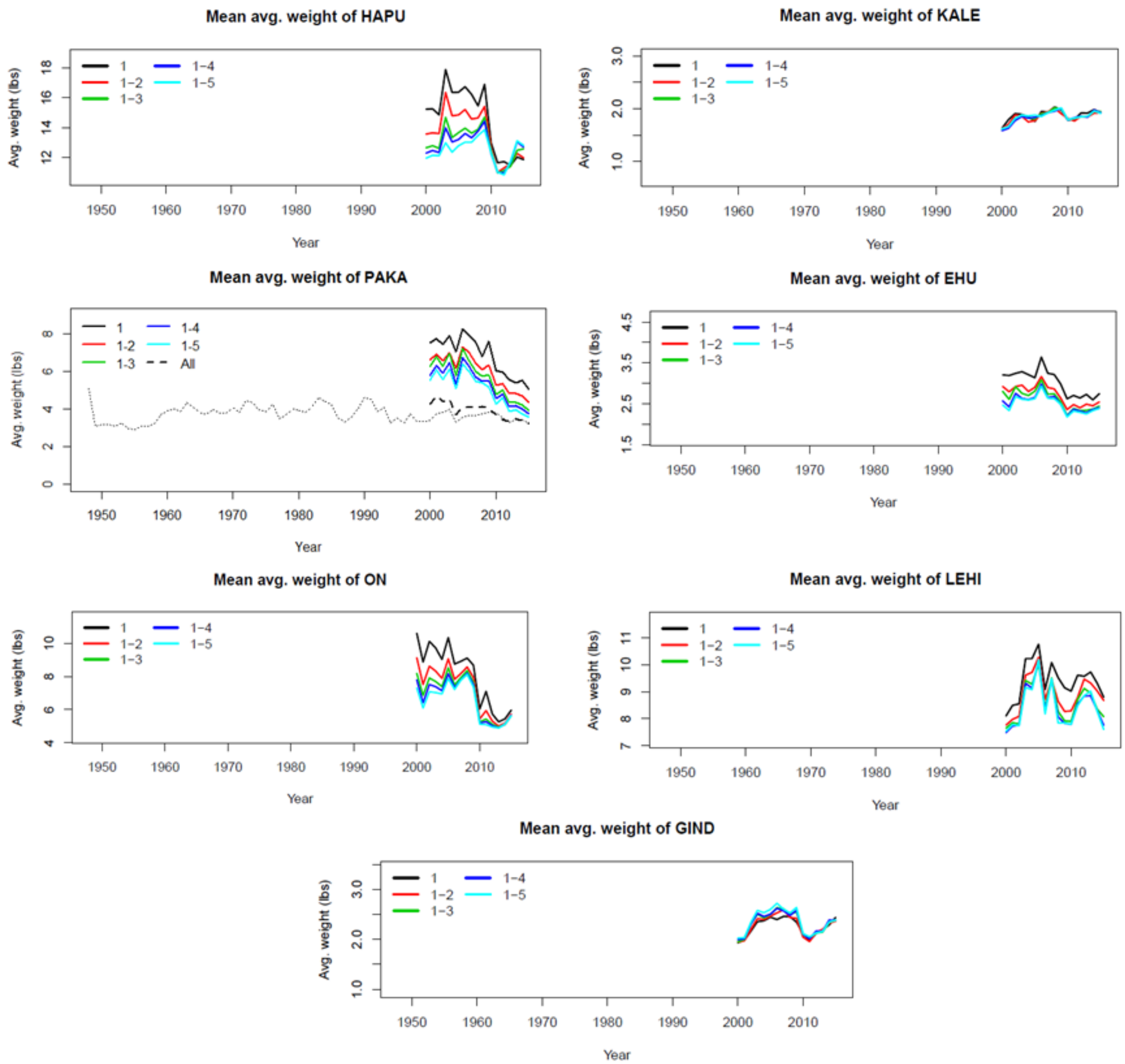


Figure 5. Individual species average annual weights as calculated from the DRS data set, for various combinations of fish number per record. In the Opakapaka figure, the dashed line represents average annual weights from all records and the dotted line from 1948 to 2015 represents the average weight from all records reporting Opakapaka in the FRS data.

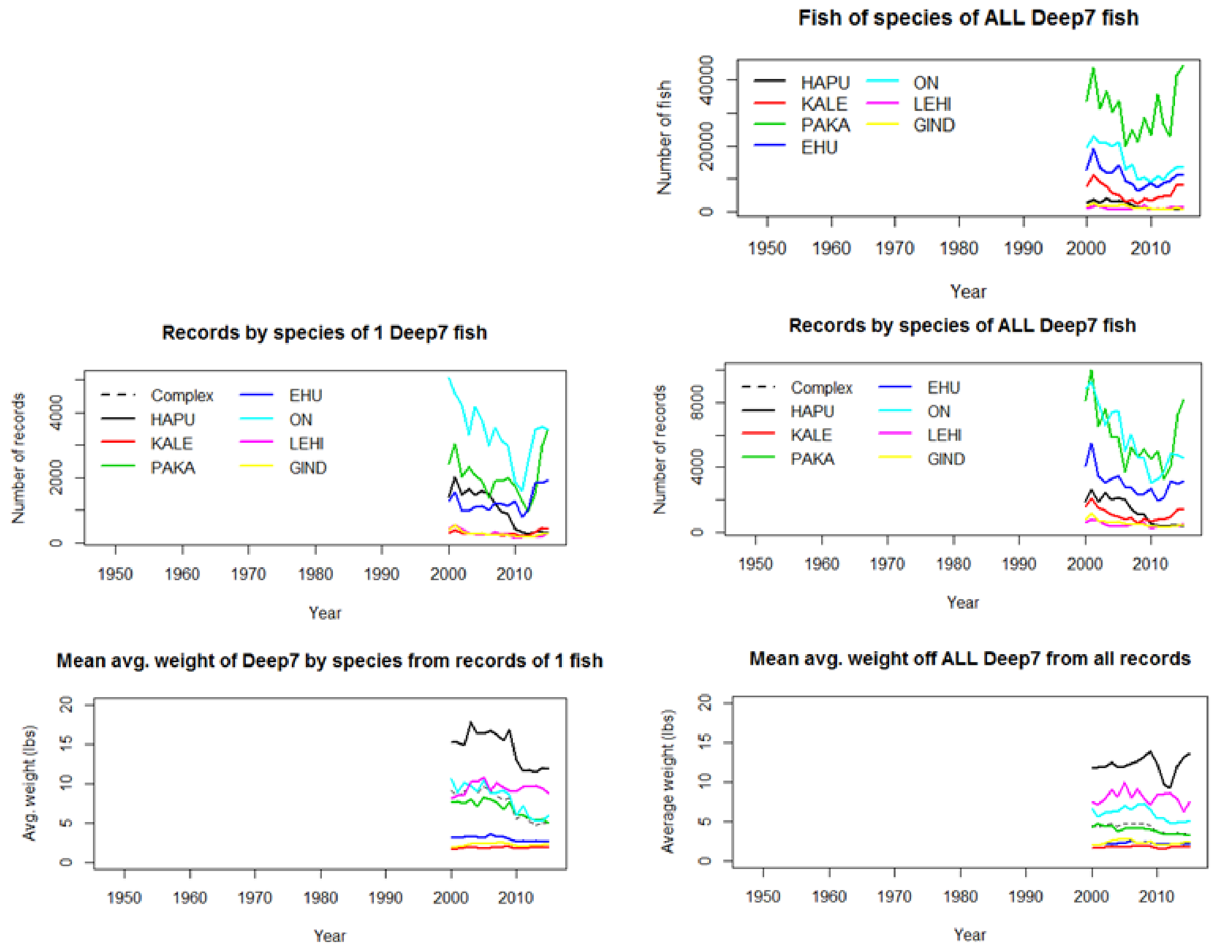


Figure 6. (Top right) Number of fish in the DRS data by species. (Middle left) Number of records reporting 1-fish weights, and (middle right) reporting weights in the DRS data set by species and for the Deep 7 complex. (Bottom left) Average annual weight from DRS records reporting 1-fish weights, and (bottom right) reporting weights by species and for the Deep 7 complex, with the same color key as the middle figures.

D.2.2. Initial Filters to Select Representative Weight Data for this Fishery

Having resolved to use the fisher-reported FRS data set for fish weights, as well as for catch, a similar agreement was made to use FRS records for Deep 7 bottomfish species (species in Table 12) in areas in the main Hawaiian Islands (areas in Figure 1, Table 2) to represent fish sizes. All fish weight analyses from this point forward use FRS data. Fishers report the number and total weight of fish by species. If more than 1 fish of a single species is reported, average weight can be calculated by dividing the total weight by number of fish. This calculation does not capture weight variability among fish reported in the same record. The quality and quantity of fisher-reported weights of bottomfish caught varies. Data is sparse in some cases, especially for earlier years such as 1955–1975, and for certain species, such as Kalekale, Ehu, and Gindai.

A histogram of the total fish weights reported for each record including records with multiple fish, indicates that the distribution is skewed toward smaller total weights, but also that there is some tendency for reporters to round off weights as evidenced by the small spikes in number of total weights at intervals of 5 (Figure 7). No filter was pursued to address this rounded data, which seems to approximate the rest of the data distribution fairly well.

When presented with an annual time series of average weights (weighted by the number of fish for each record), it was evident there was a drastic change in these weights from state fiscal year 1948 to 1949, likely due to the fewer number of fish reported in 1948 because reporting only existed for half that year (Figure 2). Based on this, participants agreed to exclude weight records from January 1, 1948, (the start of the data set) to June 30, 1948, because the shorter 6-month duration of data is not representative.

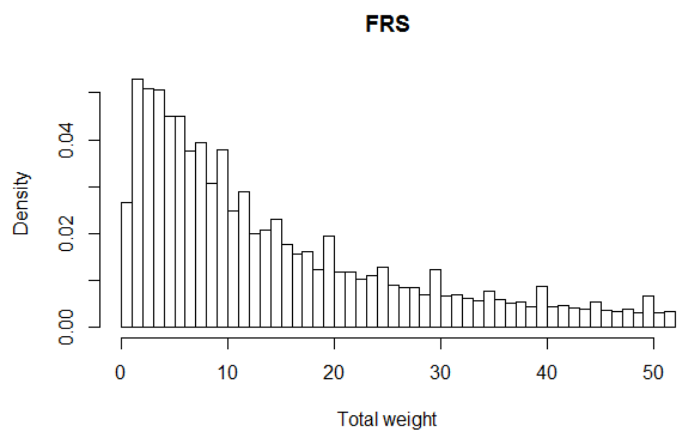


Figure 7. Frequency histogram of total weights from Deep 7 bottomfish FRS records in the main Hawaiian Islands.

Participants also discussed the utility of selecting by a specific gear type, since PIFSC scientists explained the concept of selectivity and how it is specific to a gear type. For CPUE calculation, only records that reported using bottomfishing gear (gear type = 3, deep sea handline) were used (see also Section C.3, below). A similar filter was suggested for weight data, which would result in excluding approximately 5% of the data on Deep 7 bottomfish in the main Hawaiian Islands. One fisher expressed concern that his colleagues' misreporting of gear would result in their data being excluded, but Reggie Kokubun confirmed that for recent years DAR staff have been calling fishers when a report is submitted that reports Deep 7 species but reports a gear type that is not classified as bottomfishing gear. More often than not, that report is corrected to bottomfishing gear. Given that this is the primary gear for this targeted fishery, it constitutes 95% of the weight data available, and that using only this primary gear would make selectivity more straightforward to deal with in the stock assessment, participants decided to only select records that reported using bottomfishing gear.

A proposal was made to exclude records reporting either 0 lb, or 0 numbers of fish since an average weight cannot be calculated from these records. Applying such a filter would result in excluding 19% of the remaining records on Deep 7 bottomfish in main Hawaiian Islands reporting bottomfishing gear. Given that such records are not useable for calculating average weight, participants agreed that this filter should be applied.

As a way of quality checking for extreme errors, there was discussion on whether to filter out records whose average weight is higher than a maximum and lower than a minimum. Three sets of species-specific maximum and minimum values were presented for consideration: values designated by the DAR as associated with the FRS, values designated by the DAR as associated with the DRS, and values from the PIFSC biosampling program for samples collected from 2007 to 2014 (Table 4). The PIFSC biosampling values came from sampling by PIFSC scientists, with contracted fishers, market samples, UFA measurements, and specimens collected during the bottomfish fishery-independent survey and the values were derived only using MHI and not NWHI samples. Some fishers asked to check on whether the maximum weight for a Kalekale from PIFSC biosampling was accurate, and it was likely accurate since it was listed as identified by the UFA with the next largest fish similarly sized.

Participants had the option to select any minimum and maximum values for each species that they wanted to, and were not limited to these three sets of options. Applying such a filter would not result in many records being filtered out. As an example, applying a filter to the FRS data using the max and min values from the PIFSC biosampling data would result in removal of 879 out of more than 400,000 records, with 70 records smaller than the minimum and 809 records larger than the maximum. Keeping in mind that not all fish that are caught make it to the dealer given the 1-lb minimum weight requirement at the Honolulu Fish Auction, smaller fish will show up in the FRS data than the DRS data. Participants decided to apply this filter as an error check and use the minimum values associated with the DAR FRS data (since they are smaller) and use the maximum values associated with the DAR DRS data (since they are larger – with the exception of Hapuupuu). The max weight of Hapuupuu from the FRS data was discussed as possibly coming from a larger congener species, and so was not used.

Table 4. Minimum and maximum fish weight values for Deep 7 bottomfish considered for data quality filtering purposes. DAR values were provided by DAR and not derived from raw data. Shaded columns were chosen as data quality filters. PIFSC Biosampling data covers 2007–2014 and comes from MHI samples only.

SOURCE	**DAR: Fisher-reported Data**		**DAR: Dealer-reported Data**		**PIFSC Biosampling**	
	Min (lb)	Max (lb)	Min (lb)	Max (lb)	Min (lb)	Max (lb)
Hapuupuu	0.41	131	0.5	100	0.1	49.4
Kalekale	0.1	5	0.1	6.5	0.11	18.98
Opakapaka	0.4	27	0.9	28.5	0.12	34
Ehu	0.2	10	0.2	22	0.1	17
Onaga	0.5	40	1	50	0.1	49
Lehi	0.5	35	0.5	35	0.75	23
Gindai	0.3	8	0.3	10	0.15	6

RESOLUTION for Initial filters to select representative fish size (weight) data for this fishery:

- (ii) Select all records that report a Deep 7 bottomfish species (Table 12).
- (iii) Select all records that are for areas in the main Hawaiian Islands (Figure 1, Table 2).
- (iv) Exclude records up to June 30, 1948, as data is unavailable for that full fiscal year 1948.
- (v) Select only records that reported using bottomfishing gear (gear type = 3, deep sea handline).
- (vi) Exclude records reporting either 0 lb, or 0 numbers of fish.
- (vii) Exclude records where the reported average fish weight is either greater than a species-specific maximum or less than a species-specific minimum (Table 4).

D.2.3. Selecting Data to Calculate Individual Species' Representative Weight

Participants discussed the various ways of calculating a representative weight time series for individual species. For age-structured models, individual weights of fish are best used rather than averages. But individual weights can also be obtained from entries that report just 1 fish of 1 species. There were many questions about the status of life history knowledge for all 7 species. Complete life history information is available for 2 or 3 of the species, but PIFSC scientists continue to work on the remaining species who do have some life history information available even if not complete yet. PIFSC scientists asked that the group focus on individual species—regardless of whether life history was complete yet—and indicated that they would not conduct a single-species stock assessment for an individual species unless it had robust life history information available.

Participants explored the FRS patterns of 1-fish reports, (22% of records reporting weight and number) and compared them to patterns from 1–2 fish reports (35% of records), 1–3 fish reports (45% of records), 1–4 fish reports (52%), 1–5 fish reports (58%), and all weight reports (100% of records) to explore whether they are similar.

Records reporting only 1-fish show similar ranges to all weight records, but also include an increased number of larger fish within that range. These 1-fish records generally are distributed across similar island areas as compared to all weight records (Figure 8), but the species composition of 1-fish records differs from the species composition of all weight records with a greater proportion of Hapuupuu, Lehi, and Gindai (Figure 9). Fishers commented that it made sense that Hapuupuu was often reported as a 1-fish record. If only 1-fish records were chosen to calculate average weights, then a representative sampling across space and across size ranges would occur but there would be proportionally higher Hapuupuu, Lehi, and Gindai in the records used.

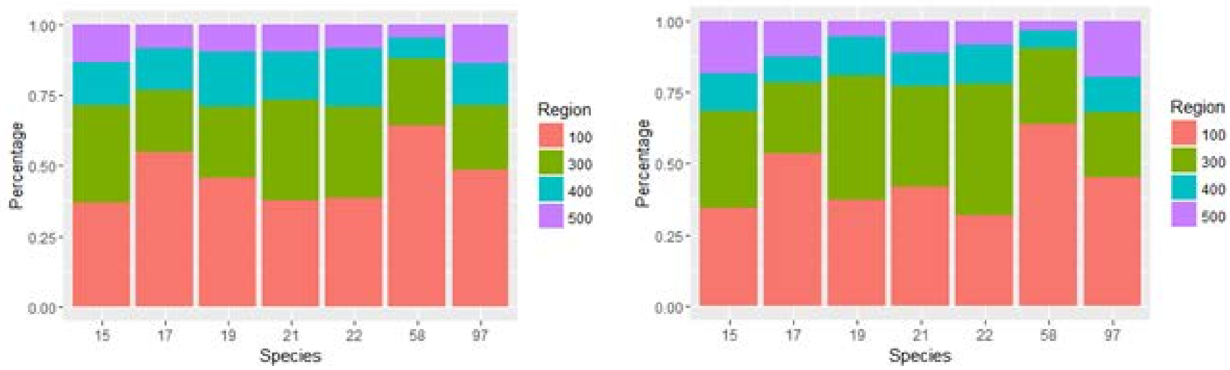


Figure 8. Spatial distribution of (left) 1-fish weight records, and (right) all fish weight records from FRS data for Deep 7 bottomfish. Pink area codes (100s) indicate the island of Hawaii, green area codes (300s) indicate islands of Maui-Nui, blue area codes (400s) indicate Oahu, and purple area codes (500s) indicate islands of Kauai and Niihau.

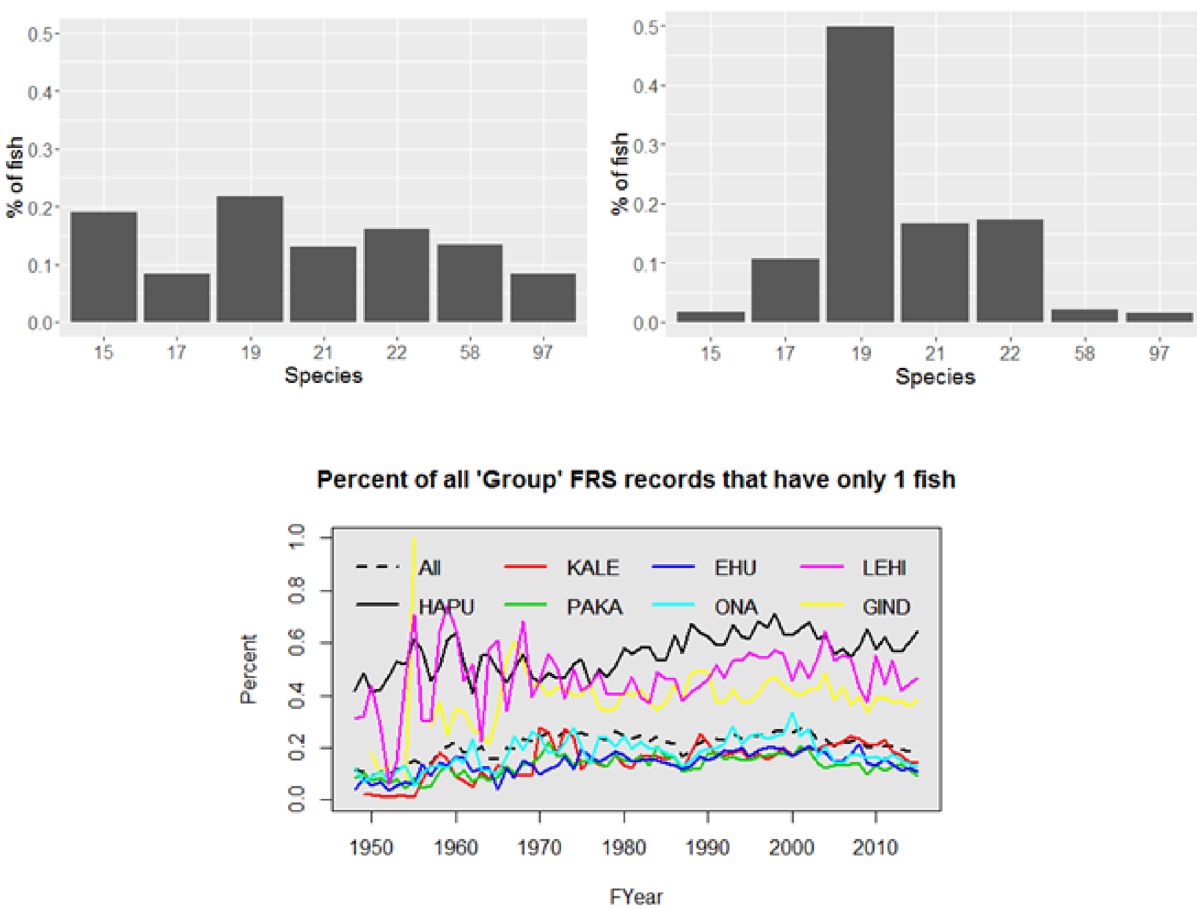


Figure 9. Species composition of FRS records reporting weights of (top left) 1-fish, and (top right) all fish. (Bottom) The proportion of all FRS records by species that are reported as 1-fish records. Species are referred to using species codes or the first few letters of their common name.

Participants examined the patterns of average annual weight for the Deep 7 complex and for individual species using different subsets of the weight data. Different subsets considered included only records reports 1-fish, records reporting 1–2 fish, 1–3 fish, 1–4 fish, 1–5 fish, 2-fish, 3-fish, 4-fish, 5-fish, 5+ fish, and all records. (Table 5) provides the percentage of weight data falling into each of these subset categories. When the values of average annual weight for each species were calculated for each of these data subsets, they differ in scale and sometimes in pattern (Figure 10). As more records are included (moving from 1-fish records only to multiple fish records), the average annual weight decreases while trends overall remain similar. Additionally, the variability in average annual weight increases if all records are included vs. if only 1-fish records are included. Participants again were dissatisfied with the idea that an average weight per record would need to be calculated unless only 1-fish records were used, because of concern that presumably the signal of fish at extreme sizes would be swamped out. A few fishers in particular were concerned that the signal of large fish would be swamped out yet they contribute disproportionately more to the reproductive output of a stock.

The reliability of the 1-fish records for calculating representative average weight was discussed in detail as a potential option for weight data. The 1-fish records were first explored because of the idea that the weights of those individual fish are known, and thus 1-fish records could be used to calculate size distributions (the opposite being true for all other weight records: averages have to be calculated for each record using the number of fish and total weight, and thus only average weight and not size distributions can be calculated reliably). Participants discussed various reasons why a fisher might report only 1 fish, including:

- The fisher only caught 1 fish, perhaps because he/she is new to the fishery.
- The fisher caught some small fish and 1 large fish, and only the large fish makes it to market and the smaller ones are eaten/retained/given away.
- The reported 1 fish is bycatch caught while targeting another species. This was noted anecdotally as happening for nighttime Onaga fisheries, that tend to catch large single Opakapaka if they catch any Opakapaka at all. Smaller single Opakapaka are caught during daytime fishing for Uku and Kalekale.
- Fishers may be more likely to catch only 1 fish for the less common species such as Hapuupuu, Gindai, and Lehi (this is corroborated by the data).
- Reporting 1 fish is a reporting artifact, especially prior to the 1970s.

Fishers noted that they generally fill in the number of fish, so ‘1’ is not a default number used. Fishers also said they do not generally separate out 1 fish from others of the same species when reporting, regardless of sizes or whether the dealer lumped into different lots by size.

Participants asked whether shifting of the scale of average weight was caused by differences in species composition, and in fact Opakapaka, Ehu, and Onaga are the 3 species which primarily show this pattern of shifting scale depending on amount of records included (Figure 11). There were some theories discussed for why these 3 species exhibit such a pattern, including schooling by size, and fishers separately reporting when a larger fish is caught. For Hapuupuu and Kalekale, the average weight from 1 fish vs. all fish records does not differ as much. The number of records and average annual weights for individual species and the Deep 7 complex combined for only 1-fish records and for all records are found in Figure 12. The trends and values of average annual weight differ by species, and as previously mentioned the scale of values and sometimes the pattern differs depending on the data subset used. Participants asked whether the average weight for the two major targeted species, Onaga and Opakapaka, change over time; results are shown in Figure 13. From 1948 to 1997, Onaga averaged 4.91 lb; from 1961 to 1997, averaged 5.13 lb; from 1999 to 2015, averaged 5.21 lb. From 1948 to 1997, Opakapaka averaged 3.77 lb; from 1961 to 1997, averaged 3.87 lb; from 1999 to2015, averaged 3.57 lb Overall, annual mean average weight varies but when averaged over a longer time period, the average weight does not vary drastically with time period.

Participants discussed the advantages and disadvantages of selecting all fish records, vs. 1-fish records, vs.5+ fish records, vs.1–5 fish records. One or two fishers thought that the average weight calculated from the 1-fish records were more representative of the average size caught by the fishery, but other fishers disagreed and found the value too high. Using 1-fish records only uses 22% of the total records in the FRS data set, and the average weight for some species was higher for 1-fish records vs. all fish records; many participants found this dissatisfactory. Some were dissatisfied with using all fish records because if more than a few fish were caught, then presumably the signal from extreme fish sizes (very large or very small) became too diluted; these persons argued to use the 1–5 fish records as a compromise, which would also allow a majority of the total records to be used. At the opposite end of the spectrum, some participants argued to use the 5+ fish records as another compromise because that time series was similar to the time series form all fish records, but presumably would be less influenced by fish of extreme sizes. Those suggesting the 5+ fish records eventually agreed that using all fish records were better than using 5+ fish records, because the former would include the larger fish seen in 1-fish records.

The decision lay between using 1–5 fish records and all fish records to calculate individual species annual mean average weight. A vote was held, and just over half of the participants voted to use all fish records. However, all but one of the participants noted that they could live with using all fish records and that many were torn between the two options. In other words, all but one of the participants noted that using all fish records was an acceptable way to calculate individual species annual mean average fish weight. Average fish weights and standard deviation are weighted by the number of fish per record.

RESOLUTION for Selecting data to calculate individual species' representative weight:

(viii) For an individual species, use all records to calculate annual average fish weights and associated error.

Table 5. The percentage of Deep 7 bottomfish records reporting weight for various combinations of fish numbers in the FRS data.

Data Subset	Percentage of Records	Percentage of Fish Numbers Reported
1 fish	22	2.3
1–2 fish	35	5
1–3 fish	45	8
1–4 fish	52	11
1-5 fish	58	14
5+ fish	48	89
All fish	100	100

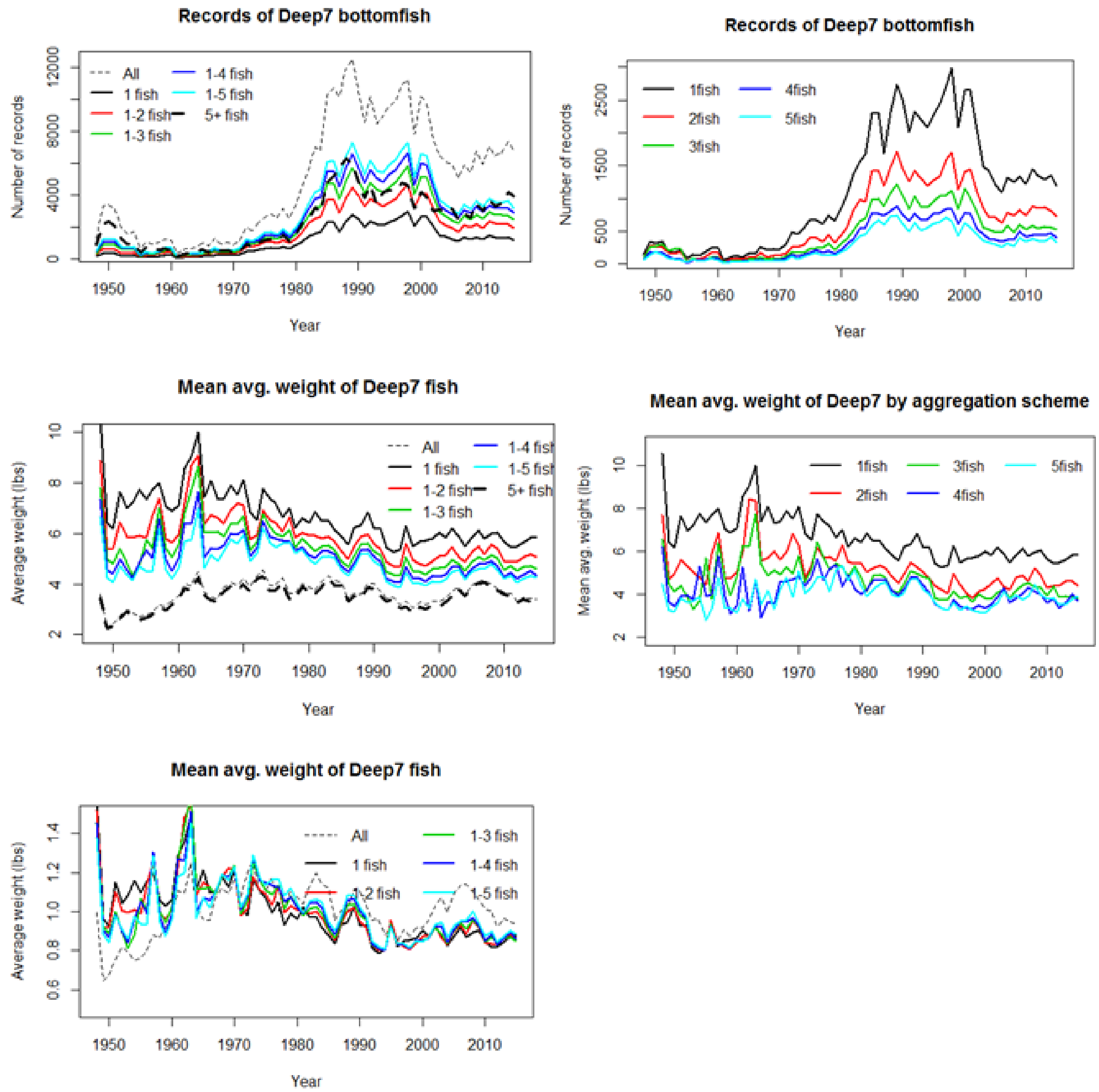


Figure 10. (Top) Number of records in FRS data set; (middle) average annual weight; and (bottom) average annual weight relative to the mean for various combinations of Deep 7 bottomfish numbers reported per record in the FRS.

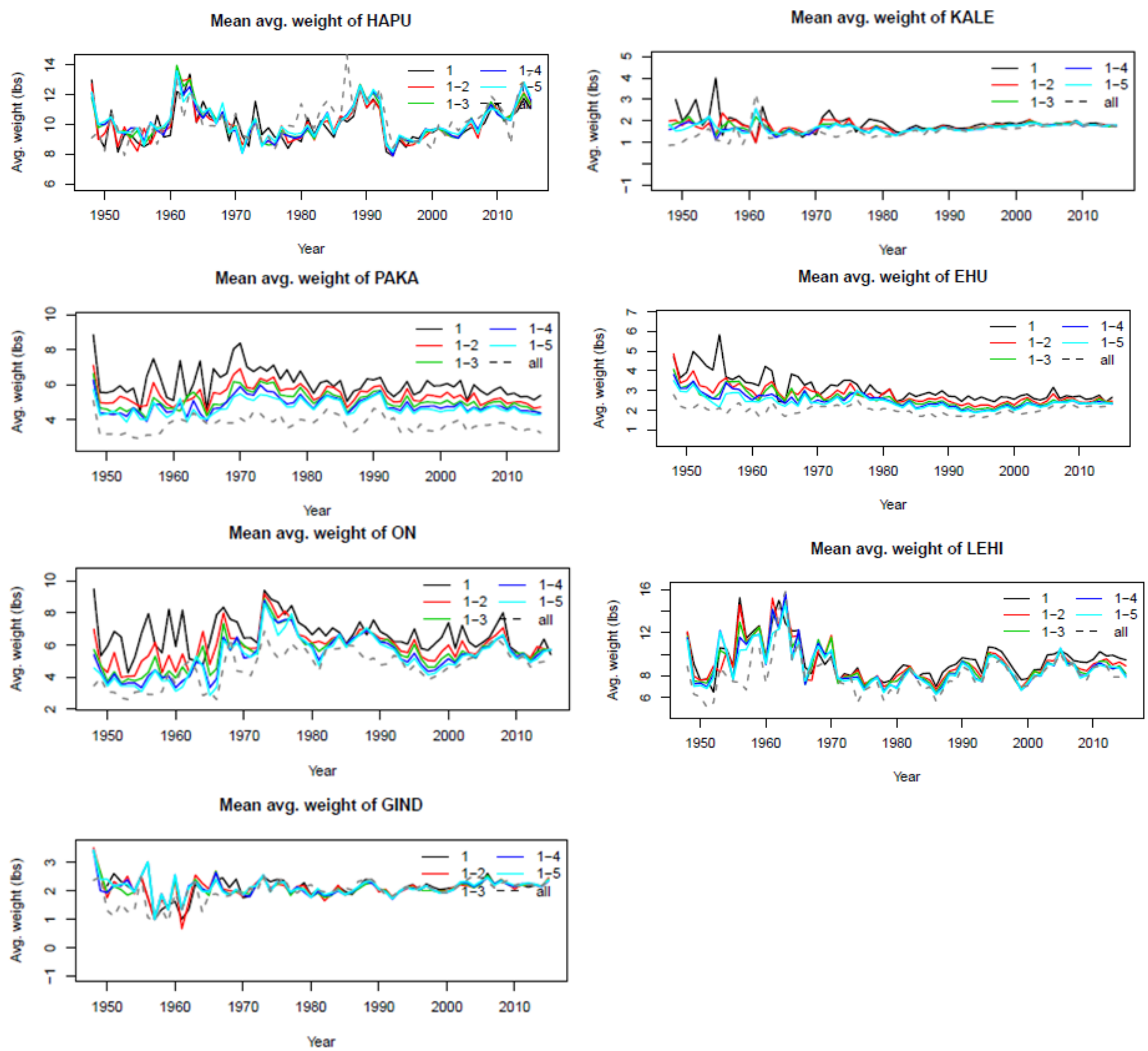


Figure 11. Average annual weights by species from the FRS data set, for various combinations of fish numbers reported per record.

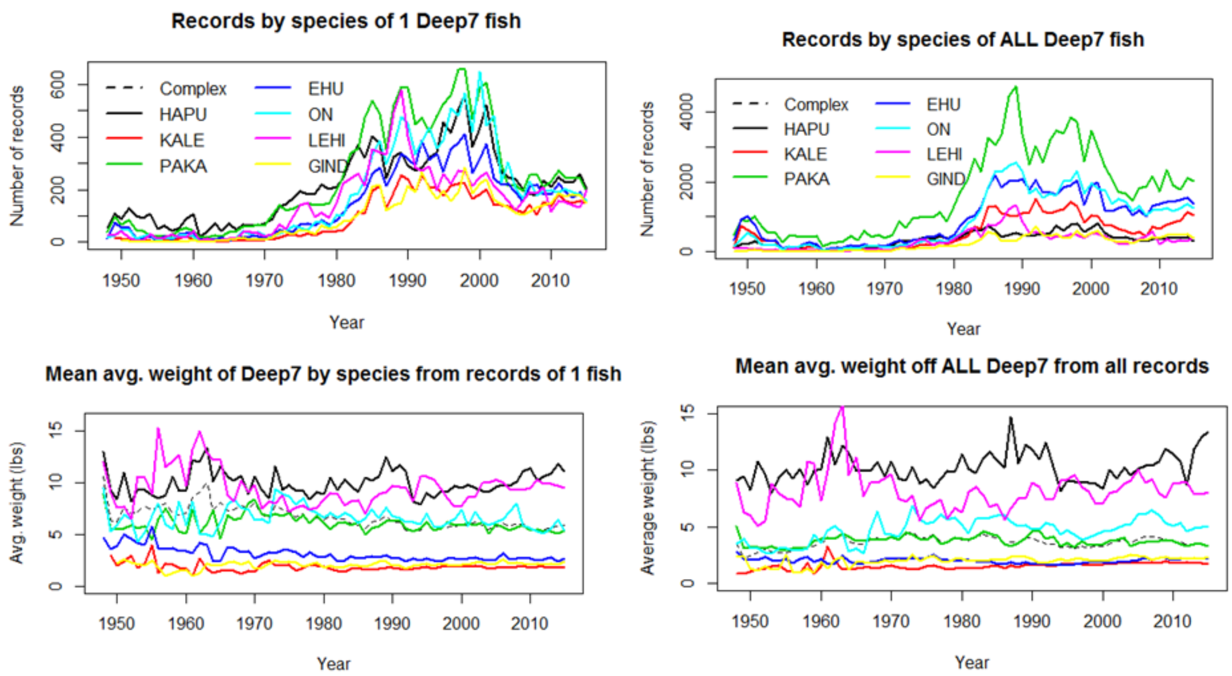


Figure 12. (Top left) Number of records reporting 1-fish weights, and (bottom left) corresponding average annual weights by species and for the Deep 7 complex from FRS data. (Top right) Number of records reporting weights, and (bottom right) corresponding average annual weights by species and for the Deep 7 complex from FRS data.

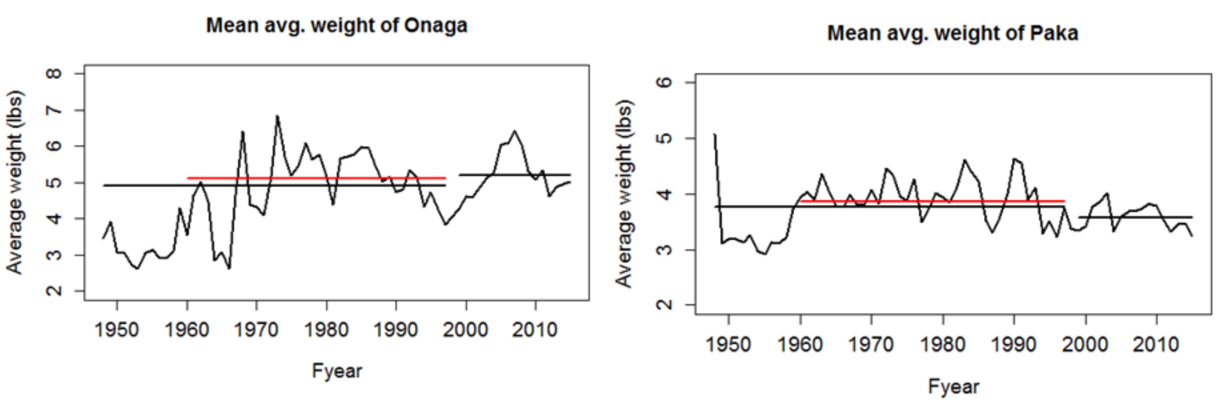


Figure 13. Mean average weight of (left) Onaga and (right) Opakapaka over time, with average weight over several time periods plotted via horizontal lines.

D.2.4. Selecting Data to Calculate Deep 7 Complex's Representative Weight

Participants discussed the challenges of coming up with a single weight metric that is representative of fish size and size trends for the entire complex, given the changing species composition over time and space and the different life histories of the Deep 7 species. If data from multiple species is combined for the complex, it's unclear how their different life history information could be combined. Several participants voiced discomfort with the idea of combining life histories. PIFSC scientists also indicated that there was not a clear path forward for combining life histories, but noted that some of the species are closer in life history than others are. Many of the relevant analyses were already discussed in the previous section on selecting weight for individual species in which participants decided to use all data available (all fish records) for a given species rather than filtering by number of fish per record; a few additional analyses are presented here.

Much of the discussion focused on the targeted species that comprise a majority of the number and weight of Deep 7 fish caught: Opakapaka, Onaga, and to a lesser extent Lehi. Opakapaka and Onaga, the primary targets, have life histories more similar to each other than some of the other Deep 7 species and that also made participants more comfortable in considering combining the two species. For example, Hapuupuu is the only grouper and has a very different life history than the other 6 snappers, and Gindai has a smaller maximum size than the other snappers.

Participants considered the mean average weight for Opakapaka and Onaga and how the values changed depending on which records to include (Figure 11). They also considered the mean average weight for all Deep 7 species combined (Figure 2, replotted in Figure 14 on a different *y*-axis scale to allow for comparison), and combinations of the mean average weight for Opakapaka and Onaga, Opakapaka and Onaga and Lehi, and Opakapaka and Onaga only on Oahu (Figure 14). The high variability in weight data in the early years for Lehi may be caused by data quality. There was a noted drop in Onaga weight around 2006, which is the year that the last set of Bottomfish Restricted Fishing Areas (BRFAs) went into effect. It was also noted that the average weight of Opakapaka as calculated from the all fish records show that the average Opakapaka caught in the fishery is larger than the size at reproductive maturity (1.55 lb for males, 2.58 lb for females; fishery average size is ~3.5 lb).

Adding Onaga to the Opakapaka data has very little effect on the overall time series, and adding Lehi to those two species has almost no effect on the time series. In fact, any time series that includes Opakapaka is similar to the time series that includes any other Deep 7 species or even all the Deep 7 species because Opakapaka constitute 50% of all fish in the data by number, and 33% of all records (Table 6).

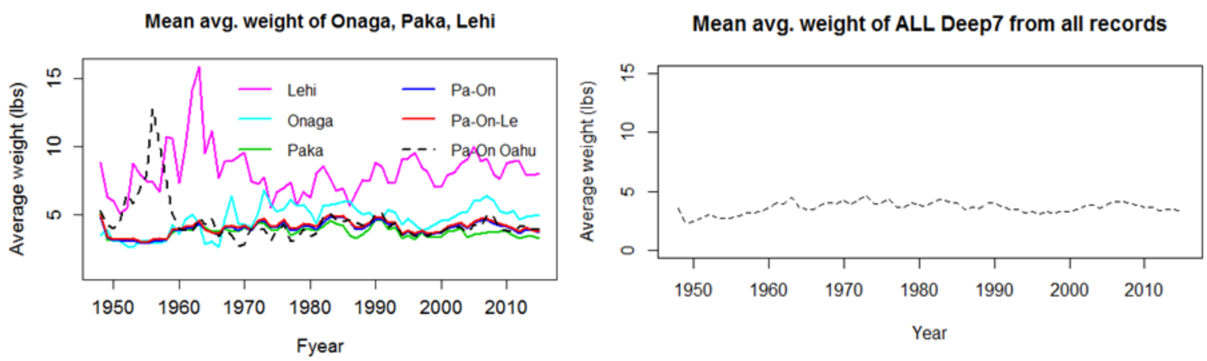


Figure 14. (Left) Mean average weight of Onaga (On), Opakapaka (Pa), and Lehi (Le) and various combinations of them including only for Oahu. (Right) Mean average weight of the Deep 7 complex.

Table 6. Percentage of the FRS data set on Deep 7 bottomfish weights by species, by number of fish and by number of records.

Species	% of Deep 7 Data by Number of Fish	% of Deep 7 Data by Number of Records
Opakapaka	50	33
Onaga	17	18
Ehu	17	19
Kale	11	11
Lehi	2	6
Hapuupuu	2	7
Gindai	2	5

The group focused on a few options brought forward by various participants as listed in Table 7. Option 1, to use all fish records for all Deep 7 species, was proposed because some participants thought it was the purest definition of the Deep 7 fishery. In other words, to calculate a weight representative of fish size in the Deep 7 fishery, all data from all species should be included. Option 2, to use all fish records for Opakapaka, was proposed because it is the dominant species caught by the fishery in terms of number, and as a result the average weight pattern dominates whenever combined with other species. Using only one species to represent the complex makes selection of life history parameters straightforward, and life history information for Opakapaka is readily available now. Option 3, all fish records for Opakapaka and Onaga combined, follows similar logic but combines the two species dominant in terms of number and targeting. Option 4, 1–5 fish records for Opakapaka and Onaga combined, uses the two dominant species but also attempts to prevent the signal from extreme fish sizes (very large or very small) from becoming too diluted by controlling the number of fish per record used to calculate average weight. Option 5, 1–5 fish records for Opakapaka, focuses on the dominant species and also seeks to prevent the signal from extreme fish sizes from becoming too diluted.

A first round of voting was held among the 17 participants present, with results shown in Table 7. Option 5 received 0 votes and was not further discussed, but the remaining 4 options received almost equal numbers of votes. Discussion on these 4 continued as participants explained their rationale for their chosen votes. It was noted that anytime Opakapaka data were included in the weight calculation, its signal dominated regardless of which other species were included. Thus, all 4 options being considered actually had similar results with the exception of Option 4, which used 1–5 fish records and thus had a slightly different trend and magnitude.

To move forward, a second round of voting was held in which participants were asked to vote for an option if it was one they could accept and live with, noting that many of the options had similar results. 14 out of 17 participants said they could live with Option 3, using all fish records from Opakapaka and Onaga combined. Still several participants seemed torn about the selection, so another vote was held to ask how many participants supported either Option 1, 2 or 3 given that they had similar results. 13 out of 17 participants said their first choice was Option 1, 2 or 3, demonstrating that the majority of participants supported using all fish records rather than 1–5 fish records. The group agreed to use Option 3, all fish records from Opakapaka and Onaga, for calculating weight information representative of fish size for the Deep 7 fishery using FRS data.

Table 7. List of options considered by participants for calculating a weight time series for the Deep 7 complex using FRS data; option in gray was chosen.

<table>
<tr><th>Options</th><th>Votes in Round 1</th><th># of Participants who Accept the Option</th><th># of Participants who Supported Option 1, 2 or 3</th></tr>
<tr><td>1. All Deep 7, all fish records</td><td>4</td><td>7</td><td rowspan="3">13 out of 17</td></tr>
<tr><td>2. Opakapaka, all fish records</td><td>5</td><td>10</td></tr>
<tr><td>3. Opakapaka and Onaga, all fish records</td><td>4</td><td>14</td></tr>
<tr><td>4. Opakapaka and Onaga, 1–5 fish records</td><td>4</td><td>8</td><td></td></tr>
<tr><td>5. Opakapaka, 1–5 fish records</td><td>0</td><td></td><td></td></tr>
</table>

RESOLUTION for Selecting data to calculate Deep 7 complex’s representative weight:

(ix) For a Deep 7 complex-level fish weight time series, use an annual average fish weight and associated error that combines all Opakapaka and Onaga records.

D.3. CATCH-PER-UNIT-EFFORT (CPUE)

A catch-per-unit-effort time series that accurately reflects the trends in CPUE over time can be standardized and used as a relative abundance index in a stock assessment.

Approach: Select effort that represents targeting of Deep 7 bottomfish, and use the corresponding catch to that effort metric calculate CPUE. Note that using a subset of records to calculate a CPUE does not interfere with using all records for calculating catch or weight information. This section is organized by steps required to complete the approach.

Issues:

D.3.1. CPUE Step 1: Selecting Data Records Considered as Bottomfish Records

Workshop participants agreed to continue using the previously-applied filter of selecting only records that report using deep-sea handline gear (gear type = 3). Bottomfishers use and report only this gear, and any records that do not report this gear type were unlikely to be bottomfishing. The prominence of this gear is evident in the data, where approximately 97% out of more than 800,000 records that reported catching Deep 7 species reported using this gear type. A number of data quality issues were discussed and addressed.

Duplicated records: Duplicated records exist in the data, defined as records for which there is a second record where every field was exactly the same. There were 663 duplicate records by only 58 unique individual fishers, 3 of which were responsible for more than two-thirds of the duplicated records. The duplicate records show up starting in 1997, with the bulk occurring between 2000 and 2010 and no duplicates after 2014. Possible reasons why duplicate records exist were discussed. It is possible, but unlikely, that a fisher catches a bottomfish in one area, moves to a different one, then comes back and catches a bottomfish of the same size in the first area. This would result in a duplicate record. Reggie Kokubun advised that the duplicates occur during a time when DAR staff would double check with fishers that these records are accurate. If accurate, if the fisher did not remember, or if no response was obtained, then the duplicate record was retained in the data but there was no note recorded to distinguish which of these 3 scenarios was involved. Given that some of these records are accurate and some may not be but there was no way to distinguish between the two, workshop participants agreed to retain duplicate records. They also noted that there are very few duplicate records compared to overall number of records (663 out of more than 800,000).

Invalid area codes: Upon investigation of the area codes defined as MHI, there are records that report invalid area codes including area = 0. Reggie Kokubun found that 323 out of 605 MHI area codes are invalid for bottomfish because they refer to fish ponds or do not correspond with a physical area, existing on paper only. A total of 1552 records that report deep-sea handline gear also reported an invalid area code, which is a small number out of 800,000 records. These are likely the result of either reporting error, or data input error, or a combination. Since there is no way to determine the correct area code, the participants agreed that any records with invalid area codes including area = 0 should be removed and not used for CPUE analysis. To implement this decision, only records that include areas on the agreed-upon MHI map (Figure 1, Table 2) will be selected.

Invalid license number = 0: Another data quality issue is that approximately 21,000 records report that license number = 0, which is an invalid number. All of these also reported the name field as “NA” so there is no information in these records that can connect to an individual fisher. Records reporting license = 0 occurred throughout the time series until 1997, after which there are no such records. Most of the license = 0 records occurred in the late 1950s, with another pulse occurring in the early 1980s. Proportionally, records with license = 0 were most prominent in 1953 when 40% of records reported license = 0; and from 1963 to 1968, ~ 20% of records reported license = 0. After 1968, records with license numbers = 0 ranged from 0 to 10% of the records, and after 1997 no records reported license = 0 (Figure 15). Since including license = 0 would result in all such records being grouped together as one fisher when the grouped records could represent any number of fishers, workshop participants agreed to remove any records with license = 0.

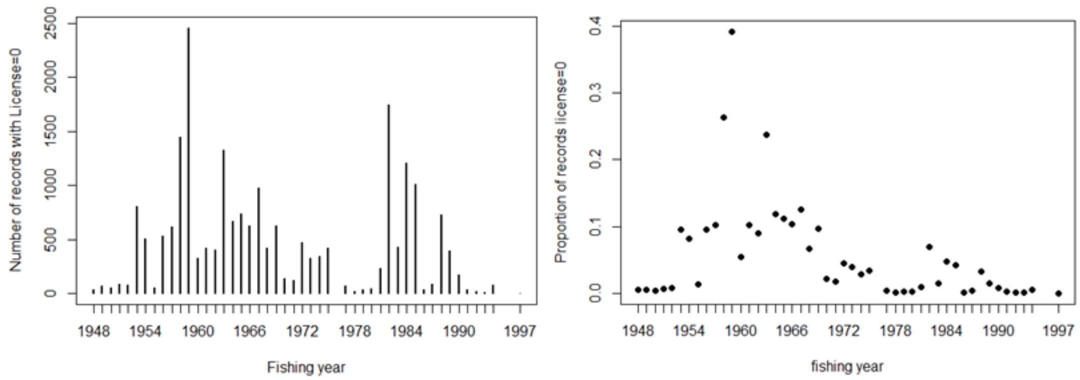


Figure 15. Number of records (left) and proportion of records (right) that report deep-sea handline gear, and also report license = 0.

Bottomfish targeting: The next topic discussed was how to select records that were targeting bottomfish. The DAR FRS data structure do not make it clear which records that caught bottomfish should be considered bottomfishing records or if another species was primarily caught or targeted. Fishers that set out to catch bottomfish may catch pelagic species while trolling to and from bottomfishing sites, and fishers that set out to catch non bottomfish species may end up catching and reporting bottomfish species. A selection method needs to be applied to the DAR FRS data to select data records considered as bottomfishing records for the purpose of calculating CPUE. There was much discussion about the best filters to use, noting that bottomfish are caught as nontarget species by some fisheries, and that bottomfishers catch other species as well. Workshop participants agreed upon several filters to exclude single-reporting days that did not target bottomfish as described in the rest of this section.

Species composition of catch is one metric that can be used to determine whether a single reporting day was targeting bottomfish. From 1948 to 2002, using the previous fisher reporting form, fishers did not have the option to record catches of 0 lb or numbers. Deep 7 bottomfish can be caught when targeting pelagic species, and when targeting Uku. Fishers who report using deep-sea handline gear but then catch mostly non-bottomfish species may be misreporting their gear type since there are multiple hand line gear type codes.

A total of ~113,000 single-reporting days made up of ~194,000 records reported using deep-sea handline gear, but did not catch any Deep 7 species. These records comprise between 20 and 70% of the records in a given year (Figure 16). Of single-reporting days that reported deep-sea handline gear starting in 1982 but did not catch Deep 7 species, 21% of the days reported Uku, 18% of the days reported Yellowfin Tuna, 10% of the days reported Papio/Ulua, and 5% of the days reported Kahala. This calculation was completed on data starting in 1982 because unique gear codes for palu ahi fishing, and for ika shibi fishing, were introduced in 1981 to describe handline gear used primarily for tuna. This provided a sense that many of these single-reporting days may not have been targeting Deep 7 species and workshop participants took the approach of removing single-reporting days that were likely to not be targeting Deep 7.

This approach involved removal of single-reporting days that reported deep-sea handline gear, caught 0 lb of bottomfish, and also either caught pelagic species, or caught Uku, or reported species = 0 which is an invalid code. The rationale for excluding based on catch of pelagic species was that these are more likely to either be trolling trips targeting pelagics such as yellowfin tuna with misreporting of gear type, or they are palu ahi fishing trips off the Big Island (but see later in this section for specific filter to address palu ahi fishing). The rationale for excluding based on catch of Uku is to remove days that definitely targeted only Uku. The rational for excluding based on species = 0 is simply that this is an invalid code.

Figure 16. (Left) Number of records that reported deep-sea handline gear, that either caught Deep 7 species or did not catch Deep 7 species. (Right) Proportion of records that reported deep-sea handling gear that also reported catching Deep 7 species.

In previous assessments, one filter used was to exclude single reporting days whose bottomfish catch was less than a certain proportion of the total catch from that single reporting day. The exact value of this proportional cutoff used in a stock assessment has changed from 90% (Moffitt et al., 2006) to 50% (Brodziak et al., 2009) to 17% in the last assessment based on a tradeoff analysis (Brodziak et al., 2011). In other words, for CPUE calculation in the most recent stock assessment, a single reporting day was excluded if catch on that day was less than 17% bottomfish. Changing the cutoff proportion changes the number of records and pounds of Deep 7 that would be excluded from the CPUE calculation (Figure 17). The exact cutoff proportion does not affect the overall trend in CPUE over time, as explored in the 2011 stock assessment (Brodziak et al., 2011).

Workshop participants further explored the data for species composition patterns because they generally liked the idea of applying a quantitative analysis. One such analysis was focused on Uku, since many fishers target Uku on the same fishing days they target Deep 7 bottomfish using the same gear to catch both. For records reporting deep-sea handline gear, ~15,000 records catch 100% Uku. Close to 600,000 records reporting deep-sea handline gear and catching Deep 7 bottomfish catch, caught 0 Uku. In fact, records from single reporting days that caught Deep 7 bottomfish, caught mostly Deep 7 bottomfish (Figure 17).

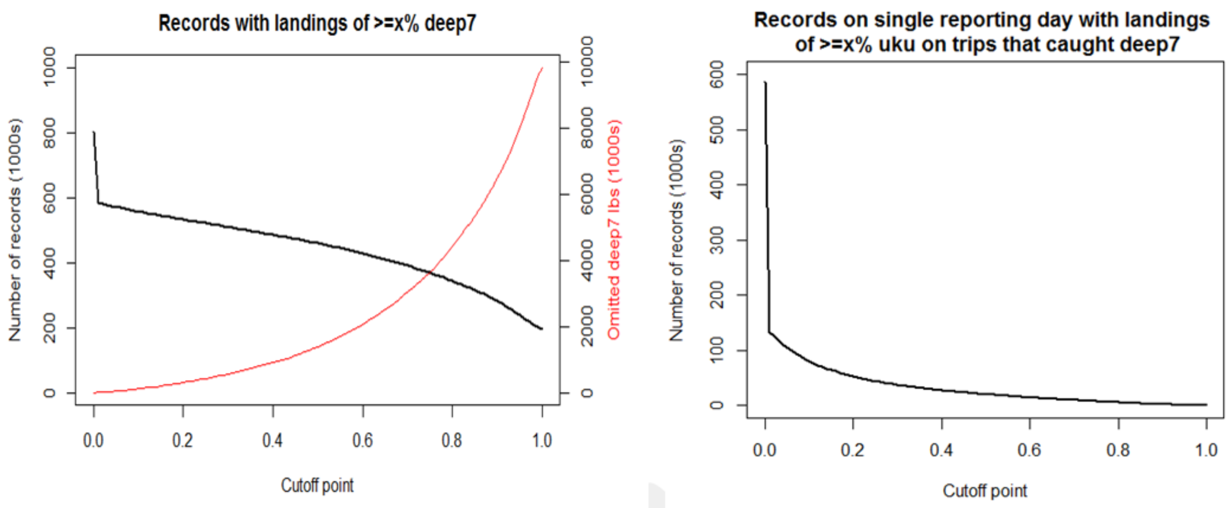


Figure 17. (Left) The number of records in the MHI that report using deep-sea handline gear, and the pounds of Deep 7 bottomfish that would be omitted, as a function of different percentage cutoff points for proportion of Deep 7 bottomfish caught in a single-reporting day. (Right) The number of records in the MHI that report using deep-sea handline gear as a function of different percentage cutoff points for proportion of Uku caught in a single-reporting day.

There was much discussion about what to do with single-reporting days that also caught Uku. Uku is an important targeted species that is caught with the same gear as Deep 7 gear, but at slightly shallower depths. Many fishers consider the bottomfish fishery to be a 3-snapper fishery (Onaga, Opaka, Uku), yet the management structure has defined Deep 7 as a management unit to be assessed. Fishers could not agree on any filter to deal with the fact that Uku is also a primary target, because they were not comfortable filtering out any single-reporting days that caught Uku but also caught Deep 7 species. In an effort to keep records that reported catching Deep 7 species, workshop participants instead recommended and agreed that the pounds of Uku caught on a single-reporting day should be included as a standardization factor in CPUE Step 4. Consideration of Uku pounds caught did not come up under original discussion of factors to consider under CPUE Step 4, although the topic can fall under “Hook competition (non-Deep 7 species caught)” which was voted 19th priority (Table 9). Given the amount of discussion about the importance of Uku to bottomfish CPUE, PIFSC scientists committed to including pounds of Uku as a standardization factor for the next stock assessment in addition to the other factors already agreed upon under CPUE Step 4.

Workshop participants agreed to omit records that occurred on dates when the Deep 7 bottomfish fishery was closed because the catch limit was reached. There was easy agreement that records reporting deep-sea handline during times of bottomfish fishery closure were very clearly not targeting Deep 7 bottomfish species. These closures occurred starting April 16, 2008, July 6, 2009, April 20, 2010, and March 12, 2011, and lasted until the end of the fishing season (August 31 of each year).

The final filter discussed to select trips targeting Deep 7 bottomfish deals with palu ahi fishing off the island of Hawaii. Palu ahi fishing consists of a slightly different handline gear setup, and because of the steep drop-offs that occur around the island of Hawaii, bottomfish are often caught during ahi fishing because this fishing occurs at the same depths as bottomfish. Fishers stated that most of the non-target bottomfish catch in the main Hawaiian Islands occurs due to palu ahi fishing which occurs primarily off Kona and South Point. From 1948 to 1981, palu ahi fishing did not have a unique gear code and was reported as gear type = 3 (deep-sea handline), but after 1981 gear codes 8 (palu ahi) and 9 (ika shibi) were implemented. After 1981, these three different fishing methods and targeting should be distinguishable based on gear code.

For data prior to 1981 around Kona and South Point, workshop participants investigated species composition to determine whether it is possible to distinguish targeted bottomfishing from palu ahi fishing trips. After agreeing on DAR area grid codes defining Kona and South Point as DAR grids 102, 122, 101, 121, 100, 120, 108, 128 (Figure 18), participants found that anywhere from 1% to 30% of the bottomfish caught in a single year comes from around Kona and South Point, with the average closer to 15%. Participants investigated the number of single-reporting days that caught both bottomfish and non-bottomfish (anywhere from ~10 to 800 annually), and the number of single-reporting days and records that caught bottomfish to get a sense of how much bottomfishing occurs around these grid areas. Given that these grid areas make up only a small proportion of the total area around the MHI, Kona and South Point represent a sizeable portion of the bottomfish fishery.

Participants then investigated species composition of single-reporting days around Kona and South Point when deep-sea handline was reported. The proportion of bottomfish compared to non-bottomfish caught varies from 30 to 85% when quantified in terms of numbers of fish, and varies from 10 to 80% when quantified in pounds of fish (Figure 19). Regardless of the units used to quantify the proportion, there is a noticeable pattern shift starting around 1985. Prior to 1985, the proportion of bottomfish caught is lower, and much more variance suggests different types of trips were being reported. After 1985, the proportion of bottomfish caught is higher and shows less variance, suggesting the type of trip being reported is more consistent in this later time period. This timing coincides with the introduction of a new gear code for palu ahi fishing in 1981, and likely reflects the fact that it took several years after that for fishers to consistently use the new codes when appropriate.

Based on these analyses, workshop participants agreed to apply a filter that removes any single-reporting day that caught less than 50 lb of Deep 7 bottomfish AND caught any Pelagic Management Unit Species (PMUS) from 1948 to 1985. The rationale is that targeted bottomfishing trips catch at least 50 lb a day, and any fishing day that results in less than 50 lb caught and catches pelagic species during this early time period were likely to be palu ahi fishing instead. Table 12 includes a list of PMUS.

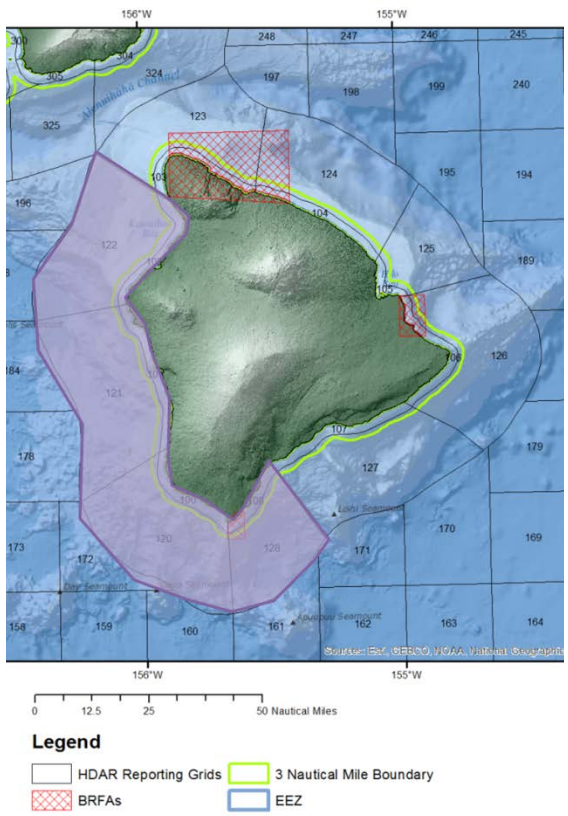


Figure 18. Grids defined as Kona and South Point around the island of Hawaii, outlined in purple.

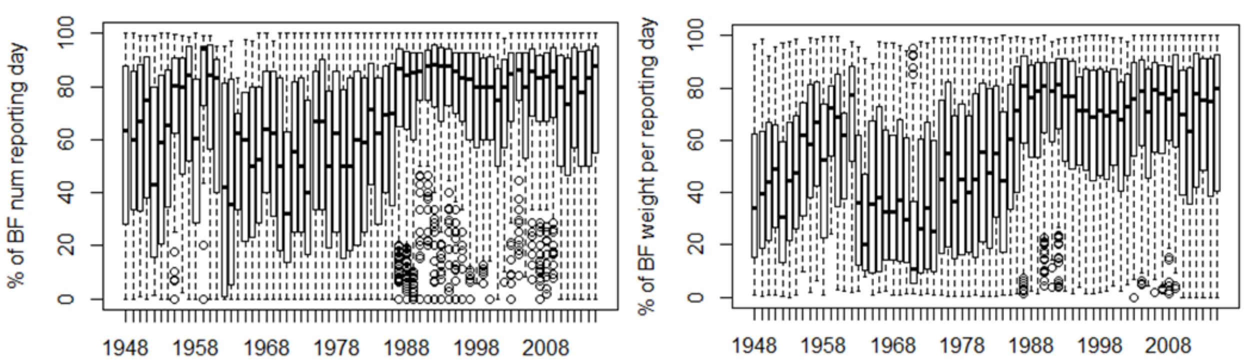


Figure 19. The percentage of bottomfish by (left) numbers of fish and (right) weight in pounds caught using deep-sea handline from Kona and South Point around the island of Hawaii. Each vertical box plot represents a single year, with annual median values as bolded horizontal bars in the center of the boxes, and the 25th and 75th percentiles represented by the bottom and top edges of the boxes.

RESOLUTION for CPUE Step 1, Selecting data records considered as bottomfish records:

(i) Select all records that are for areas in the main Hawaiian Islands (Figure 1, Table 2).

(ii) Select only records that reported using bottomfishing gear (gear type = 3, deep sea handline).
(iii) Retain duplicate records.
(iv) Remove records with license = 0.

Getting at records that target bottomfish:

(v) For trips prior to October 1, 2002, remove 'single reporting days' that report gear type 3, catch 0 pounds of Deep 7 bottomfish, and catch any of the following:
    Any pelagic species (Table 12)
    Uku (species code = 20)
    Species code = 0 (an invalid code)
(vi) Remove records that occurred on dates between when the Deep 7 bottomfish fishery was closed because the catch limit was reached and the end of the fishing season. Closures occurred from April 16, 2008, to August 31, 2008; July 6, 2009, to August 31, 2009; April 20, 2010 to August 31, 2010; and March 12, 2011 to August 31, 2011.
(vii) Remove 'single reporting days' that caught less than 50 lb of Deep 7 AND caught any pelagic species (Table 12) from 1948 to 1985 for areas around Kona and South Point (see Figure 18, grid area codes 102, 122, 101, 121, 100, 120, 108, 128) since it's likely they were palu ahi fishing rather than Deep 7 bottomfishing.

D.3.2. CPUE Step 2: Select Data Records that Comply with a Unit of Effort

Effort can be defined in various ways and its definition may even need to be changed over time based on data availability. The three major effort metrics discussed were single-reporting day, hour, and line-hour. Single-reporting day is available as an effort metric throughout the entire time series. Hour and line-hour are only available starting in October of 2002, when a new reporting form was introduced that included data fields for hour and number of lines. A few alternative effort metrics were also discussed briefly.

<u>Single-reporting day as an effort metric</u>:

A single-reporting day is defined as a unique combination of license number and date. Previously, the word "trip" was used to refer to this effort metric but that term is not accurate given the lack of information on when a fisher departed and returned to port. A single-reporting day is presumed to be no longer than a 24-hour period, but may be shorter. This effort metric is used in the absence of a unique trip identifier in the DAR FRS data. The catch corresponding to a single-reporting day is the combination of catch from all records associated with that unique combination of license number and date.

Workshop participants spent quite a bit of time discussing single-reporting day as an effort metric and noted many challenges and dissatisfactions with using this metric. It was recognized that the reported fields in the DAR FRS data prior to October 2002 do not have detailed effort metrics, and that the duration of a single-reporting day can vary from person to person and trip to trip. Ultimately participants agreed that single-reporting day will remain the effort metric to be used from 1948 to September of 2002. Workshop participants also heard from Jerry Ault and Steve Smith from the University of Miami, who were conducting a separate but complementary analysis of commercial bottomfish data. They presented several exploratory analyses they used to try and correct for multi-day trips in historical data from 1948 to 2002. Workshop participants found these analyses informative; while many analyses that Jerry and Steve used had also been independently suggested earlier, they also presented some new analyses that workshop participants had not thought of before. Ultimately the distance-based method used by Jerry and Steve suggested methods was explored as described below. The following issues with selecting single-reporting day as effort metric were discussed:

Multi-day trips need to be accounted for.

There is no way to determine absolutely whether catch reported on a single day came from that day or from multiple days of effort. This is the most prominent issue facing selection of an effort metric for data prior to 2002. Maui and Oahu historically had fishers who conducted many multi-day trips. There was a suggestion to use fisher knowledge to identify names of persons known to previously be multi-day fishers, although there is likely a limit to how extensive and how far back this knowledge goes.

In previous assessments, a **specific cutoff in terms of fish caught in a single day** was used to filter out multi-day trips. In recent assessments, that cutoff was 1500 lb of Deep 7 bottomfish caught on a single date, determined through interviews with fishers. The logic behind this daily cutoff is that this is the maximum amount of fish that a high performing fisher can catch on a single date, so reported catches higher than this amount likely occurred over multiple days but were not recorded in a way that reflects this. The maximum pounds that can be caught in a single day may differ by island, throughout time, and may be affected by market demand and price which affects how much catch fishers can and will sell on any given day.

Workshop participants investigated the implications of applying a 1500-lb Deep 7 cutoff or other daily cutoff. From 1948 to 2002, of the data that reported deep-sea handline gear,

215 (0.08%) single-reporting days made up of 1812 (0.3%) records report >1500 lb,

864 (0.3%) single-reporting days made up of 6809 (1%) records report >1000 lb, and

4672 (1.8%) single-reporting days made up of 33172 (5%) records report >500 lb.

The records are a higher percentage of total than are the reporting-days, as is consistent with more than average numbers of records for reporting days with higher than average catch. Prior to this analysis, fishers had suggested that maximum catch from a single-reporting day during the 1980s should be higher than other time periods because it was a time of maximum catch in the fishery, but in fact single-reporting days catching more than 1500 lb occur throughout the time series. It was apparent that applying this filter in past assessments removed relatively few records from CPUE analysis. Some participants asked whether catches of more than 1500 lb were catching mostly Hapuupuu as the largest fish, but in fact these catches are 29% Opakapaka, 20% Onaga, 15% Hapuupuu, 9% Ehu, and 8% Kahala by weight.

Applying different daily cutoffs changes the average weight of Deep 7 caught:

If a 1500-lb cutoff is applied, average catch on a single-reporting day is 55 lb, If a 1000-lb cutoff is applied, average catch on a single-reporting day is 52 lb,

If a 500-lb cutoff is applied, average catch on a single-reporting day is 43 lb.

Some of the fishers commented that given the composition of the fishery (a few highliners with many more being opportunistic), the ~50-lb average daily catch seemed right. Participants found use of a pounds daily cutoff unsatisfactory in dealing with the multi-day trip issue because they recounted their own experiences of highly successful bottomfishing days. Fishers noted that it was possible to catch 2000 lb of Hapuupuu on one day if a fisher found a spawning aggregation, and others recounted catching up to 900 lb. Applying a daily cutoff filter would remove feasibly accurate data on the high end of possible CPUE values, so instead other approaches were sought to account for multi-day trips.

Another investigated approach was to only use **middle reporting days**: single-reporting days when a fisher also reported catches the day immediately before and immediately after. The logic behind this exploration was that by using the date in the middle of two other reported days, there is higher certainty of selecting only a single day's effort. Only 6.7% of single-reporting days (20,442 out of 303,354) were middle reporting days, and of these 10 had Deep 7 catches > 1500 lb and 57 had Deep 7 catches > 1000 lb. The species composition of middle reporting days was investigated to determine whether they represented anomalous days. Yellowfin tuna represented 14%, Opakapa 13%, Onaga 7%, Ehu 7%, Uku 6%, Kalekale 6%, Kahala 6%, and Papio/Ulua 5% of the middle reporting days' catch by weight. Ultimately the low number of middle reporting days and the question of whether they were anomalous compared to the average bottomfishing day led the participants to reject the use of middle reporting days only.

The prevalence of **reporting only on the first and/or last day of the month** was also investigated as a means of accounting for multi-day trips. The participants asked whether fishers were prone to reporting on the first and/or last day of the month given the monthly reporting requirement between 1948 and 2002, with the idea that someone who reported on the first and/or last day of the month may have been reporting multiple trips all at once. A slight increase of records occurs on the first and last day of the month, and also an increase of records in winter months especially leading up to the end of the year because of high holiday demand for bottomfish (Figure 20). Records from the first and last days of the month did not seem to report more fish than other records: only 0.2% (167) single reporting days from the first and last days of the month caught more than 1500 lb of Deep 7 bottomfish, which is higher than the 0.08% of all single reporting days that caught over 1500 lb of Deep 7, but remains a very low number.

Another way of investigating this topic was to calculate the proportion of trips by a fisher in a given year (unique combination of license number and year; referred to as fisher-year) that reported on the first or last day of the month. The majority of fisher-years did not report on the first and last day of the month at all, but 816 (3.6%) fisher-years only ever reported on either the first or last day of the month (Figure 20). It is not known how many unique persons this number represents, because license numbers changed from year to year during the time period in question. These 816 fisher-year combinations reported 1885 single reporting days, representing 5672 records or 2.2% of the records between 1948 and September 2002. Of these 816 fisher-years that only ever reported on the first or last day of the month, 0.7% (13) single-reporting days caught more than 1500 lb of Deep 7, and 1.7% (32) single reporting days caught more than

1000 lb. The majority (512 or 63%) of these 816 fisher-year combinations fished only one time per year, 96 (12%) fished two times per year, 53 (6.5%) fished three times per year, 30 (3.7%) fished four times per year, and 32 (3.9%) fished five times per year. The most number of times fished in a year by these 816 fisher-year combinations was 12 times in a year (4 out of 816). The 816 fisher-year combinations did report higher pounds of Deep 7 caught (on average 92 lb per single reporting day, compared to 57 lb caught per single reporting day for all records), and also reported the maximum amount caught in a single-day.

Because fishers did not want to omit anyone who only fished once per year around the holidays as there are many people who fish part-time sporadically but may have been doing so for a long time, participants investigated the question of whether the 512 out of 816 fisher-year combinations that only report fishing once a year were reporting at the end of the year. However, the analysis revealed that these 512 fisher-year combinations were reporting throughout the year (Figure 20). Based on all the analyses, but primarily because the average pounds caught was higher and there are so few records in question, participants decided to exclude the data from the 816 fisher-year combinations (5672 records) that only ever reported on the first and/or last day of the month when using single reporting day as the effort metric for CPUE.

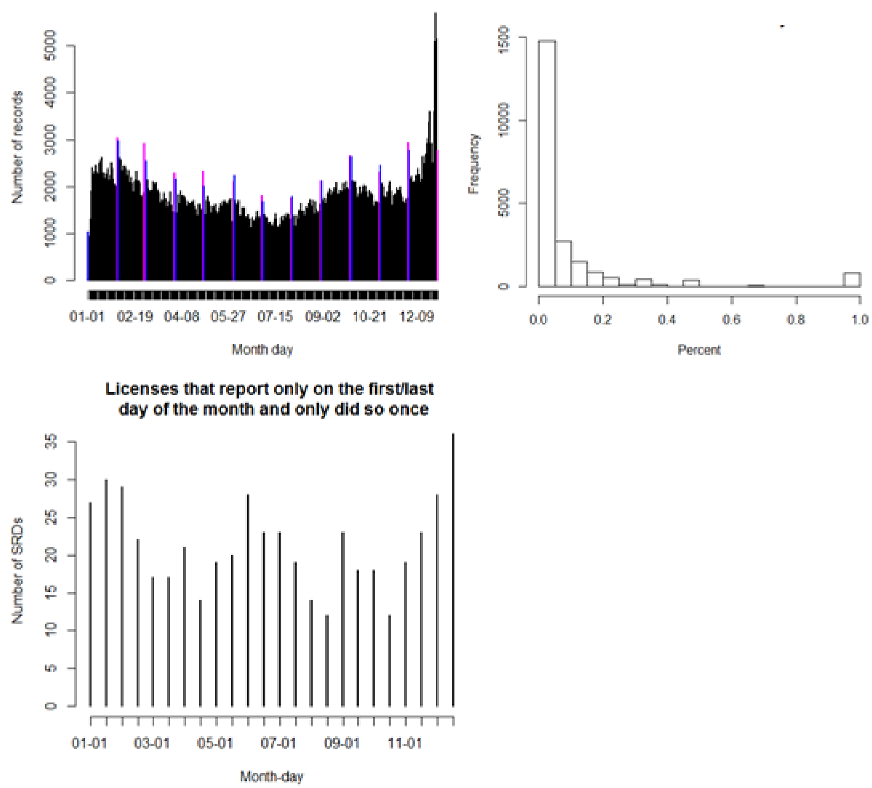


Figure 20. (Top left) The number of records reported on a given day of the year, from 1948 to September 2002. Red bars indicate last days of the month, blue bars indicate first days of the month. (Top right) The proportion of fisher-years (unique combination of license number and year) that reported on the first or last day of the month. (Bottom left) The months of reporting for 512 fisher-years who report only once a year on the first and/or last day of the month.

Workshop participants also saw presentations by Jerry Ault and Steve Smith (University of Miami) who showed some CPUE data filtering approaches that were similar, but some that were different as well. One approach discussed by workshop participants and also presented by Ault and Smith was the idea of **separating a single-reporting day into multiple days based on distances**. If a fisher reported multiple grids and/or multiple ports on a single date, and the grids are far enough apart, the trip likely took place over more than one day. For records that report both port and area fished, the distance between these two gives a sense of whether the fishing took place over one day or multiple days. Ault and Smith presented a distance of 30 nm as a threshold for the maximum distance a fisher would travel for a single day of fishing, based on previous conversations they had with fishers and the fact that the maximum speed of sampans (large vessels used for multiday trips) was 5 nm/hour, with the last sampan retired in 1996). There was much discussion by fishers that larger vessels used for multi-day trips generally traveled slower than smaller vessels, and that travel speeds changed over time based on engine capability and affordability. The practice and frequency of multi-day trips may have shifted over time, and many fishers today still fish over multiple days or at least overnight. Fishers also noted that distance is an imperfect filtering criteria, because there are fishers that travel shorter than 30 nm but stay out for multiple days. One-day fishing trips are still common practice and many single reporting days travel shorter than 30 nm as well.

Workshop participants agreed to incorporate the following two filters as a way of accounting for multi-day trips. Workshop participants recognized that these filters are imperfect in accounting for multi-day trips, and may result in multi-day trips being misclassified as single-reporting days. However, the idea is that these two filters will account for those trips most likely to be multi-day trips and leave the rest of the data alone because of lack of information to distinguish otherwise.

The first step in accounting for multi-day trips was based on whether multiple areas and/or multiple ports were reported. A few single-reporting days in the data reported multiple areas fished and multiple ports landed. Fishing in multiple areas and landing in multiple ports in one day is possible, but for some combinations of areas and ports that are very far from one another it is highly improbable and most likely represent records from multiple trips recorded together on a single day. To initially account for multi-day trips, information on the number and locations of ports and areas recorded were first used to refine the records into multiple trips. The criteria for splitting records within a single-reporting day into multiple days differed based on the number of areas and common ports visited. Many records report multiple areas visited and a maximum of 2 common ports within any one single-reporting day. All single-reporting days with one area and one common port reported were left as-is because it was assumed that these records were accurate (but a filter based on distance may be applied in the next steps).

For single-reporting days that report one area and two common ports, records were separated into two single-reporting days only when the common ports were on two separate and nearby islands (if the common ports were on the same island, the single-reporting day was left as-is). Similarly, if the common ports were at least two islands away (either Big Island-Oahu, or Kauai-Maui Nui), it was assumed to be a database error, and therefore the trip was left as-is because there was no way to know the true trip details.

For single-reporting days with multiple areas and a single common port, the distance for the trip was taken as the greatest distance among each port-area distance. For single-reporting days with multiple areas and multiple common ports, the data showed that one area was consistently associated with one common port and thus it was assumed that these single-reporting days actually took place over multiple days. These single-reporting days were separate into multiple reporting days based on the number of common ports reported. The combination of all of these filters related to separating multi-port, multi-area single reporting days into multiple reporting days, added 57 new single-reporting days to the data set.

The second step to accounting for multi-day trips was based on distance. Once distance was assigned to each single-reporting day based on the common ports and fishing areas visited, and multi-area or multi-port single-reporting days were separated into multiple reporting days based on the description above, an expected number of days were assigned to each single-reporting day based on a selected distance threshold value. After the conclusion of the workshop, as agreed there was an email discussion of whether a 30-nm threshold was appropriate or whether a different value should be used. Many fishers expressed concern that distances traveled varied by island, over time with engine capacity, and that distance traveled is not always a good predictor of how long a trip is. As examples, fishers on Maui only need to travel ~ 8 nm to reach fishing grounds but may still stay out for multiple days, and when sampan boats were still in operation, they commonly stayed out for multiple days. When the distances traveled by each single reporting day were plotted in frequency histograms by decade, it was evident that most distances traveled were between 5 and 30 nm with another cutoff around 50 nm. This pattern is stronger after 1970 (Figure 21). Based on conversations with participants at the workshops, it was expected that the threshold in the earlier part of the time series would be smaller than that in the later parts of the time series due to the vessels participating in the fishery early on being larger and slower. However, in the end it was agreed that a threshold of 30 nm would be applied to all years as it indicated a clear break in the number of trips occurring in years after 1960, and was inclusive of possible trips in the 15–30 nm range for years prior to 1960 where a clear breakpoint was slightly less obvious. Selecting a 30 nm distance was thus a conservative distance that left many of the single-reporting days as-is. Each trip was assumed to last a day for every multiple of this threshold. Thus, a distance between 0 and 30 nm would reflect a 1-day trip, a distance of 30.1 to 60 nm a two day trip, and so on. Based on this criterion, a total of 23,258 trips were changed from single-reporting days to multi-reporting days, with the longest trip being 11 days. Applying the 30-nm threshold to the two distance-based filters, most of the single-reporting days remain 1 fishing day, with few being broken into 2- or 3-fishing days based on distance traveled. There were 1.3 % of the single-reporting days that did not provide distance information, and those would remain as single reporting days for CPUE analysis. The average trip length and the idea that most trips are single reporting days is similar to present-day data when there is more detailed effort information available (Table 8). In the end, a majority of participants agreed to use the 30-nm threshold based on the fact that it was a maximum distance traveled in one day and the small proportion of corrections it would apply, with a few fishers remaining skeptical that it was the correct value. Applying this threshold does not correct for all multiple-day trips and notably does not correct for multiple-day trips that traveled less than 30 nm, since there is no information in the data set to distinguish it as a multiple-day trip. Workshop participants could not think of other options to remove multiple-day trips.

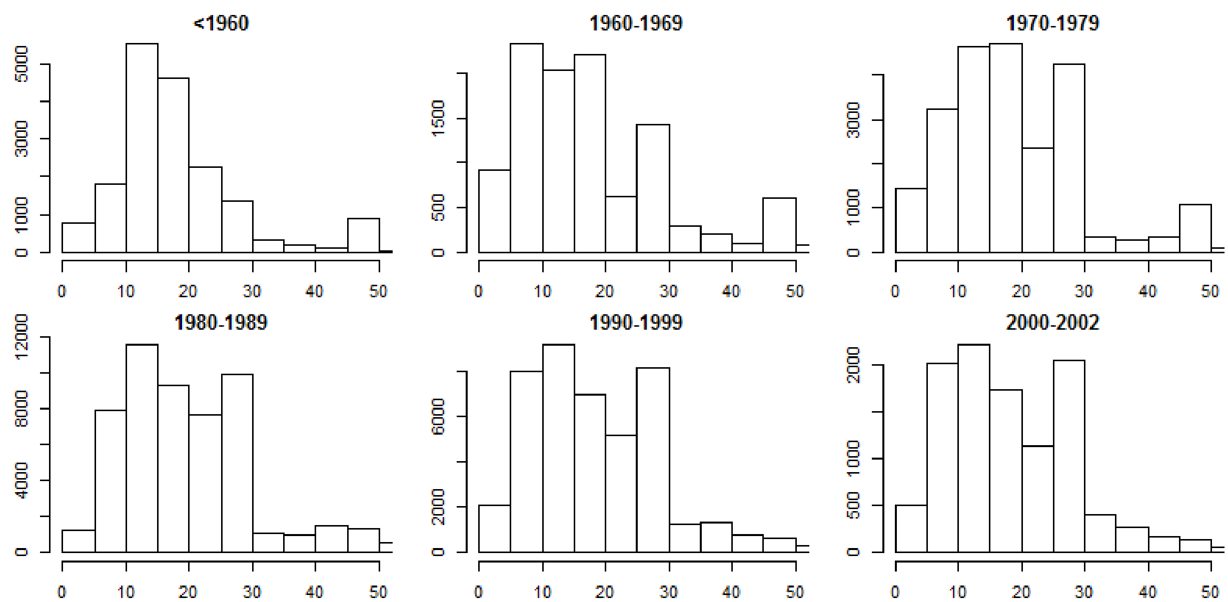


Figure 21. Frequency histograms of distances traveled by a single reporting day ('trip') as reported from 1948 to 2002, grouped by decade. There are single reporting days that traveled greater than 50 nm, but they are not plotted to maintain confidentiality.

Table 8. The breakdown of 'trip' duration if a 30-nm threshold is applied to data from 1948 to 2002, and for the most recent 2016 fishing year with detailed effort information.

'Trip' Length (Days)	Applying 30-nm Threshold to Data from 1948 to 2002		2016 Fishing-year Data
	Number of Trips	% of Trips	% of Trips
1	144292	85	86.4
2	17205	10	10.9
3	3142	2	2
4	1416	0.8	0.3
5	1495	0.9	0.4

<u>Hours as an effort metric</u>:

Starting in October 2002, fishers began reporting more detailed effort information including hours spent fishing.

Workshop participants encouraged the use of effort metrics more detailed than single reporting day in future stock assessments. Since detailed effort data are only available starting in October 2002, it was recognized that using different effort metrics over time requires splitting up the CPUE time series in the stock assessment model into before and after October 2002.

There was much discussion about the best way to report hours fished and whether to include time spent transiting, searching, or even dropping lines. There is an already-existing definition for the "hours" field agreed upon 15 years ago by PIFSC, DAR, and fishers. The best effort metric is probably total time spent fishing on fishing grounds: search time plus time lines are in the water.

This does not include transit time to and from port or between sites. Most but not all fishers report including both transit time and search time, so there are differences among commercial fishers about how hours fished is calculated. Fishers noted that it requires more time to find fish that are willing to bite than to catch once those fish are found. Regardless of how fishers have been calculating hours, if they have been reporting in the same way since 2002 then their individual trend is accurate for CPUE calculation and standardization.

Fishers also noted that there are differences in how fishers report overnight bottomfishing trips which occur over two calendar days. Some split the effort and catch over those two calendar days, and some include all on one calendar day because the effort is less than 24 hours of fishing. After October 2002, when the reporting form changed, the instructions for reporting an overnight trip of greater than 24 hours are to split the catch into 2 days but to retain as a single calendar day of fishing if the total hours fished is less than 24 hours. However, fishers report that this is still not being done correctly by all fishers. Some confusion may have arisen because of differences between reporting instructions provided via the online vs. paper reporting forms.

Workshop participants discussed and investigated a variety of different effort metrics for the time period starting in October 2002. To investigate, participants used data from 48,704 single-reporting days from October 2002 – 2015 time period. Almost all (98.8%, 48105 single-reporting days) report a single area, and 98.1% (47778 single reporting days) report a single area with a single estimate of hours fished (as opposed to different hours fished within the same single-reporting days). Number of gears (which is the number of lines for bottomfishing) was also commonly filled (47,707 single reporting days), with the overwhelming majority reporting either 1 or 2 lines fished. Other data fields were less often completed. The data field Number of Hooks is rarely completed, with only 3.4 % (1639) of single-reporting days from 25 unique license holders reporting a non-zero value. Number of Hooks does not seem to be reported since 2012. Number of hooks is only present on the bottomfishing report form, not the general report form, so it's possible that few people use the specific form that asks for number of hooks. The data field Number of Sets had no data to analyze.

Further investigation of hours fished revealed some patterns. The average number of reported hours per single-reporting day varies slightly over time but is on average ~ 7 hours. The average number of reported hours per single-reporting day also exhibits a seasonal trend, with ~ 7.5 hours fished on average in winter months and ~ 6.7 hours for summer months. Fishers reported on average more hours fished in certain areas and from certain ports/islands compared to others.

Participants compared nominal CPUE for the two primary metrics, hours fished and line-hours fished (calculated by multiplying hours fished by number of lines), with single-reporting day which is used in the previous time period. When these effort metrics were compared with the single-day fished effort metric, the overall trends in nominal CPUE were very similar (Figure 22). Overall, the selection of an effort metric did not have a major effect on resulting CPUE calculation.

Workshop participants had concerns with using line-hours as an effort metric primarily because fishers do not increase the number of lines to increase effort. Instead, they will increase the time spent fishing and searching. Additionally, fishers may not deploy all reported lines during the entire hours fished since they may drop one line if testing the waters for fish. The number of reported lines is mostly 1 or 2, suggesting that fishers are reporting this field correctly.

Other potential effort metrics were discussed. Number of Crew is a new data field reported starting in September of 2011, when fishers can now report number of crew and their license numbers. Since there are only a few years of data for this, this effort metric was not explored at this time but could be explored in the future when more years of data are available. Vessel size may affect effort, and may be more important prior to 2002 when large sampan vessels still operated. After 2002, vessel size variation is smaller, and vessel size could be partially accounted for through tracking individuals through time in CPUE standardization. Vessel-size information is available through DOBOR (Hawaii Division of Boating and Ocean Recreation) databases but they were not accessible at the time these workshops are held, nor were they accessible to stock assessment scientists in the past. Until the DOBOR database is made available, there is no information to pursue this as an effort metric. Participants recognized that this was potentially an important database and encouraged DAR and PIFSC to continue pursuing the availability of this data set. Two other effort metrics discussed but lacking any data are anchoring practices and number of line drops, the latter of which was also recognized as difficult to report accurately, especially during productive fishing trips.

Given the information presented, workshop participants agreed to use hours fished as the effort metric starting in October 2002, and did not identify other issues with using this effort metric that needed to be explored or addressed.

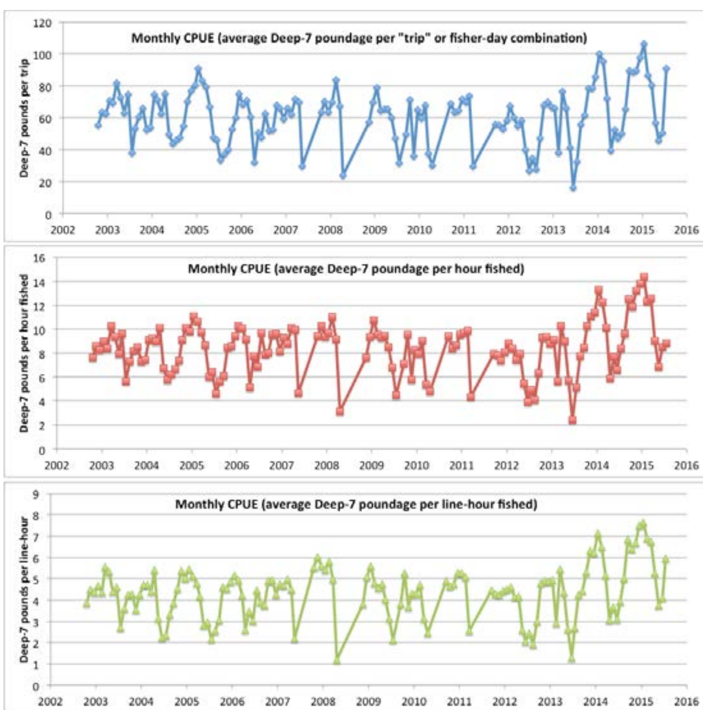


Figure 22. Nominal CPUE trends for three different effort metrics for the October 2002–present-day time period. These CPUE trends are exploratory and do not incorporate all filters agreed upon by workshop participants.

RESOLUTION for CPUE Step 2, Selecting data records that comply with a unit of effort:

From 1948 to Sep 2002:

(viii) Use 'single-reporting day' as an effort metric from 1948 to Sep 2002, defined as a unique combination of license number and date.

Account for multi-day trips showing up as a single day in this data by applying the following:

(ix) For 'single-reporting days' that report one area and two ports, divide into 2 fishing days if the two ports are on nearby islands.

(x) For 'single reporting days' that report multiple areas and multiple ports, separate into multiple fishing days based on number of ports.

(xi) For all of the above 'single-reporting days' that report both area and port, divide into 2 fishing days if distance from area to port is between 30.1-60 nm; divide into 3 fishing days if distance from area to port is between 60.1 and 90 nm; and so on.

(xii) For records prior to October 1, 2002, remove all records from licenses that only ever reported on the first and/or last day of the month.

From Oct 2002 to present:

(xiii) Use hours fished as an effort metric starting from Oct 2002.

D.3.3. CPUE Step 3: Select Data that are Representative of CPUE Trends over Time

When used as an indicator in a stock assessment, CPUE trends through time should be representative of actual changes in CPUE rather than changes in reporting or other artefacts.

Workshop participants investigated a few criteria that they agreed make a representative or responsible reporter for calculating CPUE. At the Workshop #2, workshop participants discussed the following ideas in coming to agreement on criteria for representative or responsible reporters. This list captures suggestions from discussion and does not represent consensus or decisions.

- Use the "full-time" fishers: Defined in terms of economics and not catch amount when defining the term "full-time." Possible criteria:
  - Someone who provides a report at least 9 months of the year
  - Someone who catches a minimum pounds per year
  - Someone who conducts a minimum effort per year, in terms of number or proportion of total records
- Use the "Highliners": Fishers who report most of the catch. Possible criteria:
  - Select the top 10 fishers in terms of annual catch, including all islands if this method is used because each island has a different amount of average catch
  - Select fishers from each island and calculate CPUE by island where fish was caught
  - Select fishers that average more than 200 lb on a day fished
  - Select fishers that have provided reports over a certain number of years (not necessarily consecutively), as far back in time series as possible given that fishers can only be tracked using CML as far back as 1994
- Use "mid-liners" or "low-liner" as an alternative to "highliners," given that highliners may continue to find fish even if they decline in abundance
  - Rank fishers by annual catch and select using a low enough threshold to include the mid-liner fishers, or low-liner fishers, and compare trends with highliners
- Use consistent or quality reporters: Fishers that always report the same way and/or pay attention to detail when reporting
  - Select fishers with internally consistent forms: few obvious errors, no out of range reports
  - Select fishers that submit complete forms, filling out all the fields consistently (detail-oriented reporters that may for example indicate gear change or changes in gear configuration by day fished)
  - Select fishers that report more than just target species or commercially valuable species including fish lost and released, as required in the instructions
  - Select fishers who always report on time (this filter could only be applied starting in 1994, when report date is recorded)
  - Select fishers who fished for many years (not necessarily consecutively), as far back in time series as possible given that fishers can only be tracked using CML as far back as 1994
  - Select fishers who remain in the fishery for a long time and thus show commitment or dependency on the fishery, as far back in time series as possible given that fishers can only be tracked using CML as far back as 1994

- Select fishers who do not under-report and otherwise report consistently when compared to dealer reports (this filter could only be applied starting from 2000 when dealer reports started)
- Eliminate fishers who don't report often enough to calculate a trend, for example someone who only a handful of reports a year
- Select fishers who report consistency in the duration of day fished (this filter could only be applied starting in October 2002 when more detailed effort reporting became required)
- Select fishers who are fishing regularly, reporting a minimum number of days fished a year, or a minimum proportion of the total days fished a year
- Select fishers that report zero catch records, as they take reporting seriously (this filter could only be applied starting in 1989 when it became possible to enter zero catch records)

There was much discussion on whether the use of data from "highliners" only was an effective way to calculate CPUE. A few of the scientists pointed out that the most effective fishers are also likely to be able to continue finding fish even when abundance begins declining; this ability has been shown in other fisheries. Using data from highliners only would result in a CPUE time series that would not accurately reflect trends in abundance or biomass, as highliner expertise may mask any declines. Fishers recognized that there are only a few highliners in the fishery, and using only highliners would result in ignoring the rest of the data, including data from some fishers present in the room because they were not highliners yet had been active in the fishery for decades. Not wanting to omit data from fishers who had been active over long time periods but may not have caught enough fish to be considered highliners, the group started to shift focus to embrace the concepts of including long-term or consistent fishers, since data from such persons would still be valid for purposes of CPUE calculation. This approach would be more inclusive in terms of types of fishers. Ultimately, the group decided to explore whether representative fishers could be selected based on three concepts: consistent or long-term effort, longevity in the fishery, and quality of data being reported.

It was widely recognized that the ability to track individuals through time would greatly benefit CPUE analysis and standardization. Tracking individuals is challenging prior to 1994, when individuals were assigned a different license number each year and identifying information such as first and last name were not always entered. To that end, Paul Tao at PIFSC conducted an extensive effort to add last names and vessel names all the way back to the beginning of the data set in 1948. This required taking annual lists of fisher name, license number, and vessel, and populating missing fields using some common unique field(s). In the end, tens of thousands of records annually were populated with fisher name and/or vessel name (Figure 23). Six license numbers were corrected and 12,744 more records had a last name but not a first name. Once last names and vessel names were matched, this new data set was used to explore CPUE analyses during the workshops and will be used for the stock assessment.

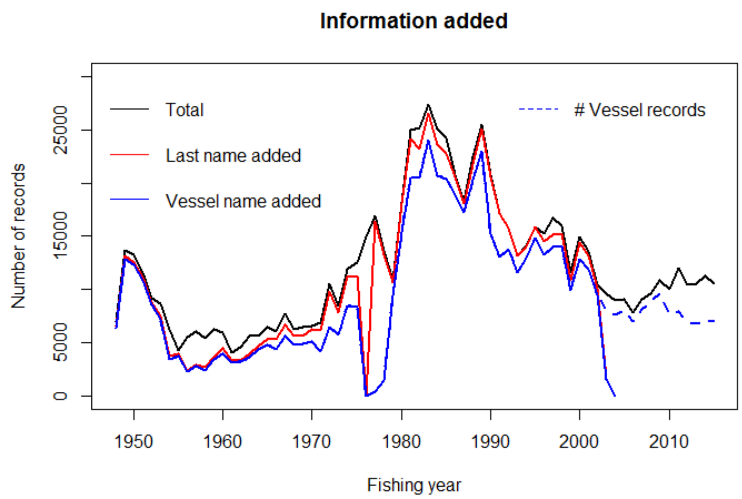


Figure 23. Number of records with information added: last name, vessel name, vessel records.

The first attempt to select representative CPUE data involved investigating whether selecting representative fishers could be based on consistent or long-term **effort**. The idea was that fishers who report many days fished a year, or a large percentage of all days fished in a year, are consistently providing quality information about the fish and fishery. Analyses of fishers with consistent effort were conducted using metrics of the percentage of available days fished a year (available days are defined as days for which at least one person reported a bottomfish). With the exception of the late 1950s and early 1960s, most years had close to 365 available days (Figure 24). Increasing the cutoff percentage of available days fished resulted in fewer fishers reporting fewer records (Figures 24 and 25). If a cutoff of 5% or more of available days is applied, ~ 75 fishers are selected in recent years. At 10% or more of available days, the number of selected fishers in recent years falls drastically to ~20. This 10% cutoff corresponds approximately to fishing a little less than once a week. Moving to a cutoff of 15% or more of available days, on average ~ 7 individual fishers in recent years are selected. Fewer than 5 fishers in recent years fished at least 25% of available days, and no fishers fished at least 50% of available days in recent years.

These analyses confirmed what fishers already know: that only a few persons fish frequently, and that most license holders only fish occasionally and often less than once a week. It also became apparent that applying any filter based on consistent effort would quickly cut out many of the fishers, since so many of them fish infrequently. A few cautioned against filtering out so many fishers, as they do make up a bulk of the effort and the records. It was noted 'highliners,' arguably the fishers who fish frequently, are skilled enough that they are likely able to find and catch fish regardless of stock reduction or movement; such ability would result in their data masking actual stock trends. Workshop participants also suggested replicating these analyses with a slightly different metric of number of days fished in a year, but the results using either metric (% of available days, number of days fished) were similar because the number of available days is close to 365 in any given year.

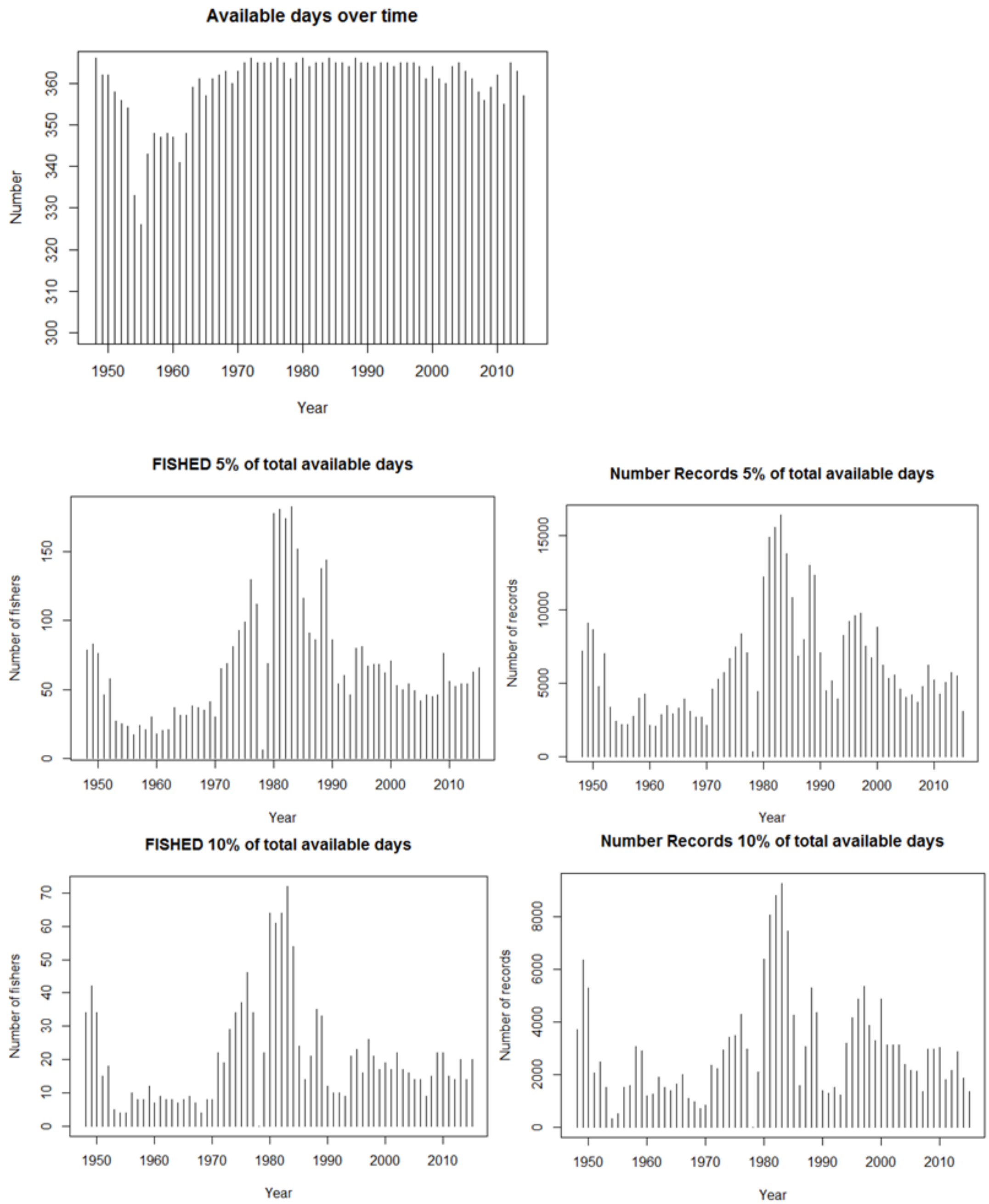


Figure 24. (Top) Number of available days (days for which at least one person reported a bottomfish) for each year. (Left) The number of fishers, and (right) the number of records for fishers that fished at least (middle) 5% of all available days in a given year, and (bottom) 10% of all available days in a given year.

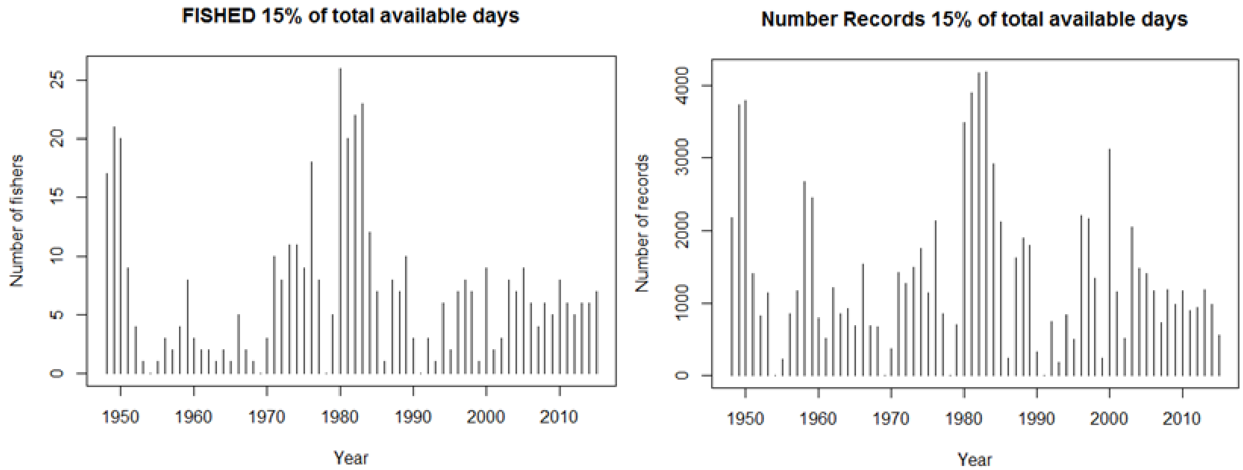


Figure 25. (Left) The number of fishers, and (right) the number of records for fishers that fished at least 15% of all available days in a given year.

The second attempt to select representative CPUE data involved investigating whether selecting representative fishers could be based on **longevity** in the fishery. The idea was that fishers who fished a certain minimum number of years (that don't need to be consecutive) are consistently providing quality information about the fish and fishery. Individuals were uniquely identified by a combination of first and last name. Of note is the inconsistency of name reporting, such that it was not possible to determine whether B. Smith, B Smith, Bob Smith, or Robert Smith, were the same or different people. For analysis purposes, these were all assumed to be different people. However, instances of adding an initial to a full name (Bob Smith and Bob J Smith) were assumed to be the same person. This decreased the unique number of fishers from 9524 to 6971 over the entire time series. According to a histogram of the number of years fished per person, 46% of all fishers only ever reported fishing in one year (Figure 26). Twenty-eight fishers have been reporting bottomfish longer than 31 years.

This confirms what fishers have always known and said, that bottomfishing is a highly specialized fishery that requires skill and expert knowledge. It is difficult for an individual to become a bottomfisher, so it seems many individuals may try it and fish for a few years but don't stay in the fishery. Very few individuals fished more than 20 years, with a handful of individuals fishing more than 30 years. Intuitively, as the number of years fished increases, the number of records reported increases as well. Workshop participants discussed the pros and cons of potentially excluding the 46% of fishers that only report in 1 year. If the fisher is reporting only in the most recent year then he/she is just entering the fishery and should not be excluded. If the fisher is reporting at some point in the past, then their individual CPUE calculation is only one point in time and does not provide a trend. The stock assessment scientists explained that these individual points from the 46% of fishers would not strongly influence CPUE calculation if individual fishers are tracked in the standardization because they do not offer trend information. In other words, these 46% of fishers would not strongly influence CPUE.

Ultimately, the workshop participants realized that effort and longevity are intertwined concepts: someone who fishes frequently is also likely to be someone who has fished for many years. Rather than exclude any data based on the analyses for effort and longevity, the workshop participants wanted to be inclusive of the bulk of the data. They agreed to remove records from the fishers who used gear 3 but never reported catching a Deep 7 species, as they did not consider these fishers to be Deep 7 bottomfishers. No other filter based on effort or longevity was chosen, but instead the participants agreed to include an additional factor for CPUE standardization under Step 4: cumulative number of Deep 7 bottomfishing trips to-date. This serves as a metric for cumulative experience. PIFSC stock assessment scientists agreed to add this to the list of standardization factors already agreed upon under CPUE Step 4.

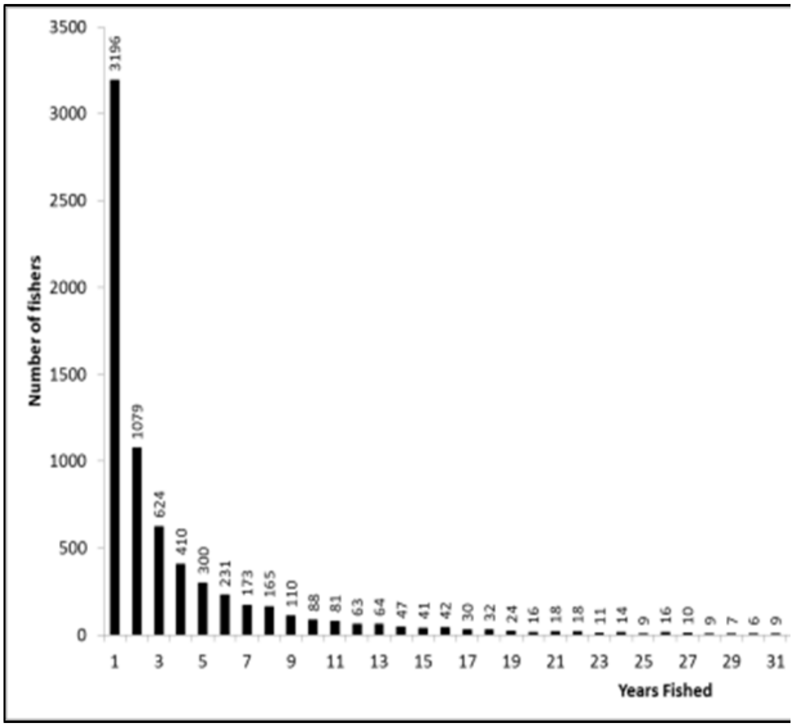


Figure 26. The number of unique fishers (tracked by name) that reported for a given number of nonconsecutive years. Twenty-eight fishers have fished more than 31 years, but their data are not plotted to maintain confidentiality.

The third attempt to select representative CPUE data involved investigating whether selecting representative fishers could be based on **quality** of data being reported by fishers. The idea was to attempt to select fishers who submitted mostly complete forms most of the time, with attention to detail and few obvious errors such as:

- No out-of-range reports for area,
- No out-of-range reports for pounds,
- Reporting in the “LOST” field,
- Reporting in the “RELEASED” field, and
- Reporting non-Deep 7 species.

There was a suggestion to investigate whether an additional criterion could be fishers who report on time as indicated by date report submitted available since 1994, but this information was not readily available in the data set so was not further explored.

Out-of-range reports for area: Given that it was impossible to know whether a reported value for most fields was an error or not, workshop participants looked at the “area error flag” field which is available starting in 2002 to get a sense of how often invalid or incorrect areas were reported. There were a total of 16 area error flags from 9 fishers, and they occurred between 2007 and 2010 (Figure 27). These 9 fishers reported a total of 2916 records starting in 1977, and the maximum percentage of records with an area error flag by any of these fishers was 6%. Based on the low number of records these 9 fishers represent, workshop participants concluded they are likely to be infrequent fishers and agreed not to remove any data based on an area error flag. It’s possible that some of these errors are filtered out already in a previous step (since the analyses shown for discussion did not include all agreed-upon steps given that they were conducted prior to agreement on all steps). Also, these area error flags are system-generated and not fisher-reported. Reggie Kokubun noted that while staff does follow up if a flag is generated, it is still possible that validated correct data still has the area error flag.

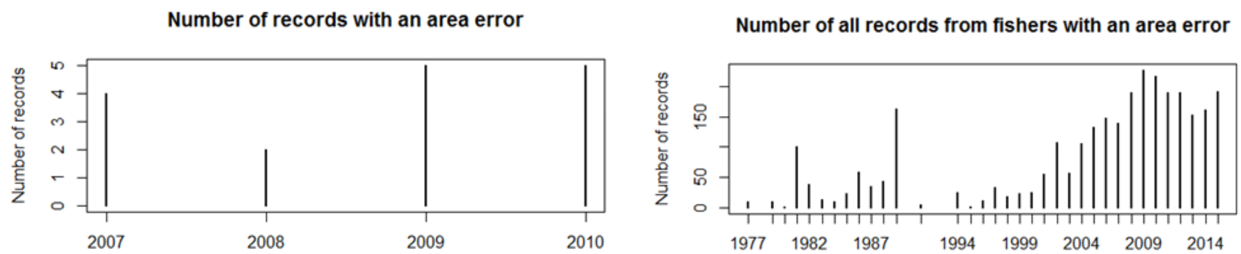


Figure 27. (Left) Number of records with an area error flag by year. (Right) The total number of records reported annually by the 9 fishers from the left panel.

Out-of-range reports for pounds: Out-of-range reporting was investigated as a possible data quality filter using the “pounds error flag” field, available starting in 2002. There were 119 records with a pounds error flag from 86 fishers who report a total of 43,301 records over the years (Figure 28). The maximum percentage of records with a pounds error flag by any of these fishers came from fishers who report less than 10 records total. Workshop participants again drew the conclusion that these are likely to be infrequent fishers based on the low number of records these fishers report, and agreed not to remove any data based on a pounds error flag. Similar to the area error flag, pounds error flags are system-generated and not fisher-reported. Reggie Kokubun noted that while staff does follow up if a flag is generated, it’s still possible that validated correct data still has the pounds error flag.

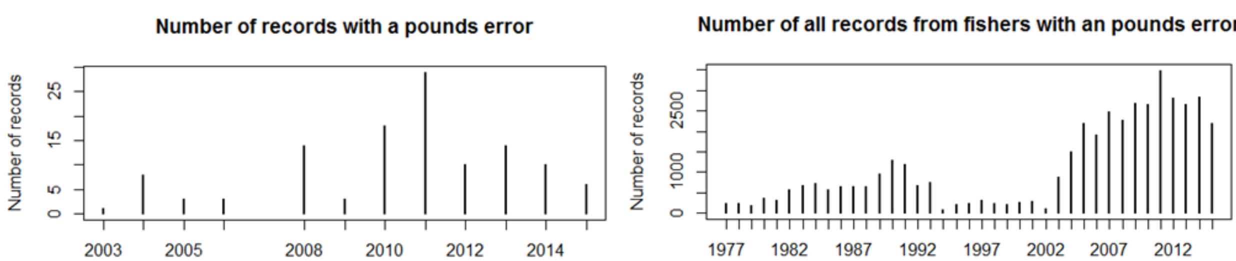


Figure 28. (Left) Number of records with a pounds error flag by year. (Right) The total number of records reported annually by the 86 fishers from the left panel.

Reporting in the “LOST” field: The “LOST” field is provided so fishers can report when a fish was lost to predation or otherwise fell off the gear. Only 8791 records reported by 691 fishers reported in this field. These 691 fishers reported a total of 152,436 records over the years (Figure 29). Workshop participants agreed there was no justification to use only these 691 fishers for CPUE calculation. Doing so would result in omitting a high percentage of the available data, and the increase in reporting in the “LOST” field over time is most likely a result of increased outreach on this topic starting in ~ 2007.

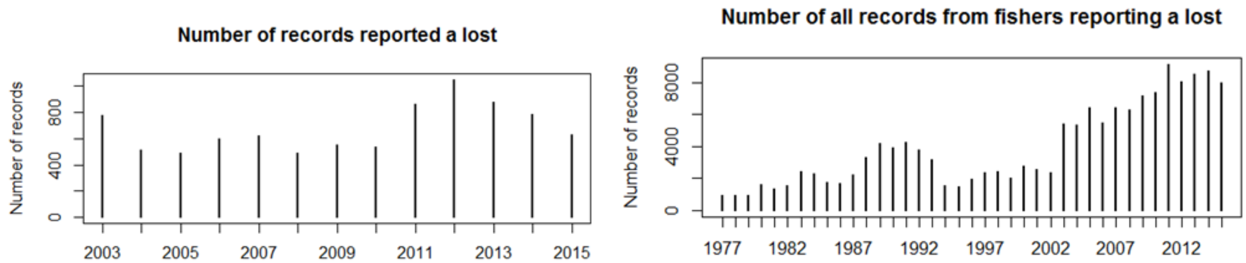


Figure 29. (Left) Number of records reporting in the "LOST" field by year. (Right) The total number of records reported annually by the 691 fishers from the left panel.

Reporting in the "RELEASED" field: The "RELEASED" field is provided so fishers can report when a fish was released back to the water. Only 6950 records reported by 508 fishers reported in this field. These 508 fishers reported a total of 105,236 records over the years (Figure 29). Workshop participants agreed there was no justification to use only these 508 fishers for CPUE calculation. Doing so would result in omitting a high percentage of the available data. Fishers commented that the only reason to release a bottomfish would be as part of a tagging study, so using only this data would not be representative of the fishery. There was a suggestion by fishers to remove data from targeted tagging trips starting around 2014 using individual fisher names and in cases where 'single reporting day's released more than 50% of Deep 7 caught by numbers for that day. PIFSC scientists agreed to investigate this for the subsequent benchmark stock assessment rather than the ongoing on.

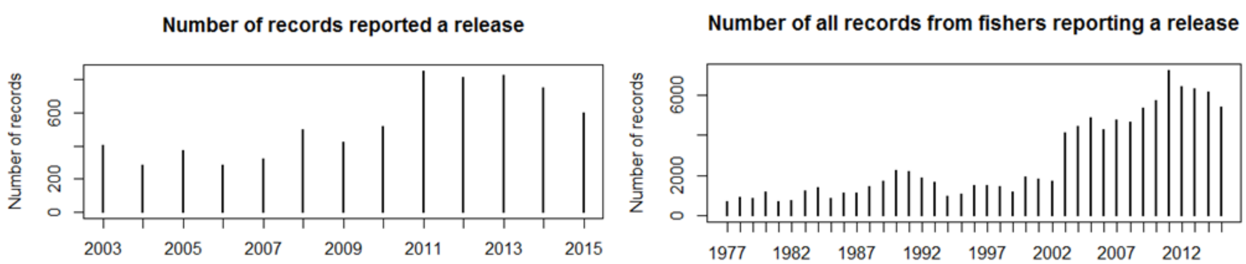


Figure 30. (Left) Number of records reporting in the "RELEASED" field by year. (Right) The total number of records reported annually by the 508 fishers from the left panel.

Reporting non-Deep 7 species: Workshop participants investigated the occurrence of fishers reporting species that were not Deep 7, as a potential method for selecting fishers that report carefully. A total of 388,290 records reported by 8870 fishers showed species that were not Deep 7. These 8,807 fishers reported a total of 388,290 records over the years (Figure 31). Nearly all the data is included if these 8,870 fishers were selected. Fishers commented that this made sense given that often non-Deep 7 species were caught in the process of targeting Deep 7 species. Workshop participants agreed not to place special emphasis on this data and thus agreed not to include a filter to use only this data.

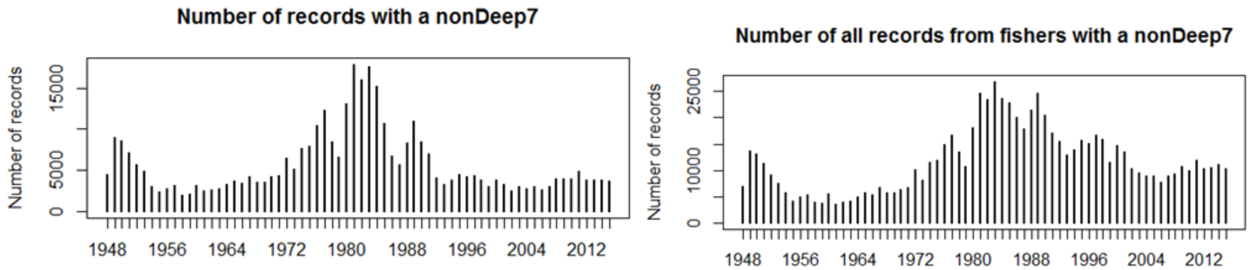


Figure 31. (Left) Number of records reporting non-Deep 7 bottomfish by year, 388,290 total by 8870 fishers. (Right) The total number of records reported annually by the 8807 fishers from the left panel.

Finally, workshop participants agreed to remove records associated with the fishery-independent survey efforts that have taken place over the most recent years. These records do not represent bottomfishing, since they take place in areas designated by PIFSC scientists and using standardized gear configurations and fishing times. There was discussion and acknowledgement that removing these records may be difficult and incomplete given that there is not a specific flag for these records, and that fishers participating in the survey did also bottomfish commercially on the same days they also fished for the survey, sometimes even in the same grid area.

RESOLUTION for CPUE Step 3, selecting data representative of CPUE trends over time:

(xiv) Remove records from fishers who used gear 3 but never caught any pounds of Deep 7.
(xv) Remove single-reporting days that occurred during fishery-independent sampling for specific fishers in specific grids.

D.3.4. CPUE Step 4: Account for factors that affect CPUE

Certain factors are known to affect CPUE that can be accounted for, provided there is data available on them. Workshop participants identified a list of factors known or believed to affect CPUE, and then identified those for which data are available and thus are feasible to explore in an analysis. Recognizing that some factors affect CPUE more than others, factors vary in their data availability, and that investigating all possible factors would require extensive time and effort, participants then prioritized analysis of the factors.

At workshop #2, participants came up with a list of factors that are known or believed to affect CPUE are feasible to investigate because of data availability. Recognizing that to investigate all of these factors would require extensive time and effort, participants prioritized these 24 factors at workshop #3 through a voting exercise to obtain equal input from all participants. Starting with the list of factors known or believed to affect CPUE and have data available, participants were asked to choose their top 5 priority factors that he/she believed most affects bottomfish CPUE. The most important factor believed to be affecting CPUE received a vote of “5,” the second most important received a vote of “4,” and so on. The numbers were tallied for each factor and those with the highest total numbers were considered to be the most important factors affecting CPUE as determined by workshop participants. The following table shows the prioritized list of factors affecting CPUE for which data is thought to be available.

Table 9. Factors that are known or believed to affect CPUE and are feasible to investigate because of data availability, in order of priority by workshop participants.

Factor	Category	Priority	Total number	Total # of votes
Fisher skill (license number as proxy)	Fishery	1	69	20
Wind speed and direction	Oceanography	2	62	19
Fishing grid area	Ecosystem	3	49	13
BRFAs and Kahoolawe	Management	4	33	11
Island area	Ecosystem	5	21	6
Month or season	Oceanography	6	18	9
Technology (data for this still pending from interviews)	Fishery	7	18	6
Targeting by species	Fishery	8	12	6
Moon phase	Oceanography	9	6	2
Depredation	Ecosystem	10	5	2
Ecosystem productivity (chlorophyll, SST)	Ecosystem	11	4	3
Fish market prices since Oct 2002	Economics	11	4	3
Fisher abundance	Fishery	13	4	1
El Nino/La Nina	Oceanography	13	4	1
TACs and ACLs starting in 2007–2008	Management	13	4	1
State noncommercial bag limit starting June 1998	Management	16	3	1
Distance to fishing area from port	Fishery	17	2	2
Pacific Decadal Oscillation	Oceanography	18	2	1
Rainfall (may affect recruitment?)	Ecosystem	19	0	0
Taape catch	Ecosystem	19	0	0
Hook competition (non-Deep 7 species caught)	Ecosystem	19	0	0
Ciguatera (effect of UFA no longer buying kahala in early 1980s)	Economics	19	0	0
Federal noncommercial bag limit starting in 2008	Management	19	0	0
Targeting by size	Fishery	19	0	0

Workshop participants present for this voting exercise self-identified as 9 scientists, 8 fishers, 3 data analysts, and 2 managers. PIFSC agreed that for the next stock assessment, the first 3 priority factors will be explored for inclusion in CPUE standardization. For each subsequent stock assessment, the next 2–3 factors on the list will be explored. There is likely to be diminishing returns on effort in exploring factors further down the priority list, since many factors received few or no votes by workshop participants. At workshop #4, fishers requested to include additional standardization factors in the upcoming stock assessment and PIFSC scientists agreed to explore two more standardization factors in addition to the first 3. The first factor is catch of Uku on the same single-reporting day, which ties into the targeting by species factor listed as priority 8 and the hook competition factor listed as priority 19. The second factor is the cumulative number of Deep 7 single-reporting days so far, which is meant as a proxy for experience and is similar to priority 1 which is fisher skill.

Although voting occurred using an agreed-upon list of factors, 2 workshop participants wrote in ‘vessel size’ as a factor and gave it a total priority of 7. Vessel size as it affects CPUE was discussed in detail under CPUE Step 2, “Other effort metrics.” It was noted that vessel size may affect effort, especially in the early period from 1948-2001 when there were more larger vessels that were more likely to conduct multiday trips and carry more crew. After 2002, vessel size variation is smaller, and vessel size is accounted for through tracking individuals through time.

Table 10 is a comprehensive list of all factors that are known or believed to affect CPUE developed at Workshop #2, regardless of whether there is data available for them. Participants also noted that a combination of these factors could affect CPUE in synergistic ways. Workshop participants recognized that currents and gear configuration are two factors that are very important in affecting CPUE, but noted there is no data set available on them so they cannot be investigated.

Table 10. Comprehensive list of all factors known or believed to affect CPUE and corresponding general category, in no particular order. Data is not available on all of these factors.

Factor	Category
Fisher abundance	Fishery
Fisher skill (license number as proxy)	Fishery
Technology (data for this still pending from interviews)	Fishery
Targeting by species	Fishery
Targeting by size	Fishery
Distance to fishing area from port	Fishery
Travel time due to weather/wind	Fishery
Behavior	Fishery
Bait type	Fishery
Palu or not	Fishery
Anchor or drift or parachute	Fishery
Line set up or method	Fishery
Vessel characteristics	Fishery
Fish abundance	Ecosystem
Fishing grid area	Ecosystem
Island area	Ecosystem
Ecosystem productivity (chlorophyll, SST)	Ecosystem
Ahi CPUE	Ecosystem
Taape catch	Ecosystem
Depredation	Ecosystem
Bait stealing	Ecosystem
Hook competition by other species (non-Deep 7 caught)	Ecosystem
Competition or interference by other species	Ecosystem
Fish behavior (the “bite,” as affected by environmental conditions)	Ecosystem

Factor	Category
Patchiness of fish distribution	Ecosystem
Rainfall (may affect recruitment?)	Ecosystem
Day or night	Ecosystem
Current speed and direction	Oceanography
Current speed and direction at surface vs. at depth	Oceanography
Tides	Oceanography
Moon phase	Oceanography
Ocean temperature at depth	Oceanography
Wind speed and direction	Oceanography
Month or season	Oceanography
Holidays	Oceanography
El Nino/La Nina	Oceanography
Pacific Decadal Oscillation	Oceanography
Fish market prices since Oct 2002	Economics
Ciguatera (effect of UFA no longer buying kahala starting in early 1980s	Economics
Gas prices	Economics
Bait prices and availability	Economics
BRFAs and Kahoolawe	Management
TACs and ACLs starting in 2007-2008	Management
State noncommercial bag limit starting in June 1998	Management
Federal noncommercial bag limit starting in 2008	Management
Weight limit	Management

E. CONCLUSION

These 5 workshops sought to find solutions to identified commercial fishery data quality issues to the extent possible. Many possible filters were explored but not adopted for various reasons ranging from exclusion of too much data to not having enough information to apply the filter or having to make too many assumptions to apply the filter. Many new filters were agreed upon that should improve the quality of the commercial data used for future stock assessments. Workshop participants agreed according to the Terms of Reference, that each issue will not be revisited unless additional information on this past data becomes available in the future. All participants provided insightful contributions about the data and the fishery that have allowed for improved commercial data sources as inputs for stock assessments.

F. REFERENCES

Brodziak, J., R. Moffitt, and G. DiNardo.
2009. Hawaiian bottomfish assessment update for 2008. Pacific Islands Fisheries Science Center, National Marine Fisheries Service, NOAA, Honolulu, HI 96822-2396. Pacific Islands Fisheries Science Center Admin. Rep. H-09-02, 93 p.

Brodziak, J., D. Courtney, L. Wagatsuma, J. O'Malley, H-H. Lee, W. Walsh, A. Andrews, R. Humphreys, and G. DiNardo.
2011. Stock assessment of the main Hawaiian Islands Deep7 bottomfish complex through 2010. U.S. Dep. Commer., NOAA Tech. Memo., NOAA-TM-NMFS-PIFSC-29, 176 p. + Appendix.

Brodziak, J., A. Yau, J. O'Malley, A. Andrews, R. Humphreys, E. DeMartini, M. Pan, M. Parke, and E. Fletcher.
2014. Stock assessment update for the main Hawaiian Islands Deep 7 bottomfish complex through 2013 with projected annual catch limits through 2016. U.S. Dep. Commer., NOAA Tech. Memo., NOAA-TM-NMFS-PIFSC-42, 61 p.

Courtney, D., and J. Brodziak.
2011. Review of unreported to reported catch ratios for bottomfish resources in the Main Hawaiian Islands. Pacific Islands Fisheries Science Center, National Marine Fisheries Service, NOAA, Honolulu, HI 96822-2396. Pacific Islands Fisheries Science Center, Internal Report IR-11-017, 45 p.

Moffitt, R., D. Kobayashi, and G. DiNardo.
2006. Status of the Hawaiian Bottomfish Stocks, 2004. Pacific Islands Fisheries Science Center, National Marine Fisheries Service, NOAA, Honolulu, HI 96822-2396. Pacific Islands Fisheries Science Center Admin. Rep. H-06-01, 50 p.

Moffitt, R., G. DiNardo, J. Brodziak, K. Kawamoto, M. Quach, M. Pan, K. Brookins, C. Tam, and M. Mitsuyatsu.
2011. Bottomfish CPUE standardization workshop proceedings, August 4-6, 2008. Pacific Islands Fisheries Science Center, National Marine Fisheries Service, NOAA, Honolulu, HI 96822-2396. Pacific Islands Fisheries Science Center, Internal Rep. IR-11-003, 134 p.

Ralston, S., and Polovina, J.
1982. A Multispecies analysis of the commercial deep-sea handline fishery in Hawaii. Fishery Bulletin, 80(3): 435-448.

Ralston, S., D. Tagami, and P. Shiota.
1986. An Analysis of Bottomfish Size Variation at the Honolulu Fish Auction. Southwest Fisheries Center Honolulu Laboratory, National Marine Fisheries Service, NOAA, Honolulu, Hawaii 96822-2396. Admin. Report, H-86-10, 14 p.

Yau, A., and R. Oram.

2016. Summary of 2013 and 2015 Main Hawai‘ian Islands Bottomfish Research Coordination Workshops. Pacific Islands Fish. Sci. Cent., Natl. Mar. Fish. Serv., NOAA, Honolulu, HI 96818. Pacific Islands Fish. Sci. Cent. Admin. Rep. H-16-04, 53 p.

APPENDIX A: Terms of Reference

TERMS OF REFERENCE

Bottomfish Commercial Fishery Data Workshops

Terms of Reference

November 10, 2015

Background:

The challenges surrounding the robustness and utility for scientific and management purposes of the Hawaiian Department of Aquatic Resources (DAR) commercial Fisher Reporting System (FRS) data have been known for some time. Starting at least in 2004, a series of stock assessment workshops were convened to assess current bottomfish and coral reef fish assessment methods and develop a plan to improve data collection and assessment methods to a point where reliable resource assessments could be obtained. The review, synthesis, and analysis of existing data were identified as a high priority by workshop participants. Since then, various meetings have been held to discuss, identify, and resolve ongoing data issues, including a commercial catch-per-unit-effort (CPUE) workshop in 2008. The DAR has made efforts at improvements in data collection and the NOAA Pacific Islands Fisheries Science Center (PIFSC) has continued analyzing and exploring the DAR FRS data set through multiple iterations of the bottomfish stock complex assessments.

Despite these ongoing efforts, this issue emerged again in December 2014, when NOAA contracted the Center for Independent Experts to conduct an independent external peer review of the 2014 Main Hawaiian Islands Deep 7 bottomfish stock assessment which was completed by PIFSC scientists. This independent external review panel had strong reservations regarding the quality of the Hawaii DAR commercial FRS data used to calculate input catch and CPUE for the stock assessment. The review panel concluded that although the methods to standardize CPUE and the assessment model were appropriate, the deficiencies in input data quality caused the stock assessment to have serious flaws that compromised its utility for management.

In response to these concerns and others, PIFSC will work with DAR, and will seek input from the NOAA Pacific Islands Regional Office (PIRO), the Western Pacific Regional Fishery Management Council (the Council), and fisher stakeholders to help understand and resolve these data issues once and for all. This will be accomplished by identifying, and resolving to the extent possible, issues in the DAR FRS data that affect the quality of data on Deep 7 bottomfish for stock assessment purposes. A series of workshops and small-group working sessions will be held to facilitate completion of these tasks, with the ultimate goal of producing a higher quality FRS data set that is deemed acceptable for stock assessment and management use by scientific, management, and user group (fisher) stakeholders. The intent is that this series of workshops will not be repeated each time a benchmark stock assessment is conducted, because a final data set has been agreed upon.

Scope:

Hawaii DAR data reporting on Deep 7 bottomfish or deep-sea handline gear in the main Hawaiian Islands (MHI):

- Commercial Fisher Reporting System (FRS) Data, 1948 – 2015
  This is the primary data set investigated, data through 2015 were used in workshop analyses.
- Dealer Reporting System (DRS) Data, 2000 – 2015
  This data set was investigated mainly to address issues pertaining to the FRS data set, data through 2015 were used in workshop analyses.

Participants: NOAA Fisheries PIFSC and PIRO, the Hawaii DAR, the Council, and (commercial) fishers/stakeholders

Goals:

1) Identify and agree upon issues that affect the quality and consistency of DAR FRS data on MHI Deep 7 bottomfish fishing.
2) Investigate and agree upon an acceptable resolution to each issue, understanding that some issues may not be fully resolvable and/or may not have a significant effect on stock assessment results.
3) Finalize a document describing the issues discussed and the agreed-upon resolution for each.
4) Apply the list of resolutions to a copy of the DAR's commercial FRS data to create a stock-assessment-suitable higher quality set of DAR FRS data on MHI Deep 7 bottomfish fishing for use by PIFSC in future stock assessments. This higher quality data set will constitute the best available scientific information for calculating total catch, catch-per-unit-effort (CPUE), and weight information of Deep 7 bottomfish species.

Workshop 1: Summer 2015

The goal of the first workshop will be to bring together participants to identify and agree upon a list of issues affecting the quality of the DAR FRS data for Deep 7 bottomfish. Participants will start with a draft list of issues and will discuss, edit, and possibly add to this list. The final agreed-upon list of issues should strive to be comprehensive to cover currently known issues regarding the quality and consistency of the bottomfish fishery data for the MHI. Ideally this list will not be added to in future workshops, since this final list will drive the subsequent schedule of workshops and intersessional work. At the first workshop, this list will be divided up and selected workshop participants will be asked to investigate these in smaller groups, with progress and results reported back at subsequent workshops.

Future workshop(s):

Future workshop(s) prior to the next benchmark assessment will be held in order to have smaller groups report back about progress made in addressing the issues they have agreed to investigate, and for the full workshop participants to discuss and agree on a path forward for each issue. The exact number and frequency of workshops is to be determined, and will depend upon workflow and resource availability. Each issue should be resolved to the extent possible, recognizing that some issues may not be resolvable and that some do not have a significant impact on stock assessment results. Participants will come to an agreement on the resolution for each issue, which will be documented in a final report and in source code created and used to apply these resolutions to a copy of the DAR FRS data. The entire process is anticipated to be completed sometime in late 2016 or early 2017.

Additional Information:

Any resolutions that will continue to change DAR FRS data will be documented, so that if necessary, these can be consistently applied in the future. A list of errors in DAR FRS data will be provided to DAR. Additionally, there should be a discussion with DAR to evaluate whether long term changes in the FRS data collection forms and processes are warranted, so that DAR can determine when and if such changes can be made at the source.

By agreeing to the resolution for each issue, participants agree that each issue has been resolved to the extent possible and agree that an issue will be revisited only if additional information about the existing DAR data relevant to that issue becomes available in the future (beyond the simple addition of new years of DAR data). Thus, every attempt will be made during this current process to seek and consider all currently available information in finding resolutions to each issue. Participants agree that the newly created higher quality set of DAR FRS data can and should be used for stock assessment and management purposes.

APPENDIX B: Acronym Table and Species Table

Table 11. Table of acronyms and definitions used in this report

ACL	Annual Catch Limit
BRFA	Bottomfish Restricted Fishing Area
CPUE	Catch-Per-Unit-Effort
DAR	Hawaii Division of Aquatic Resources
DOBOR	Hawaii Division of Boating and Ocean Recreation
DRS	DAR Dealer-Reporting System data set
FRMD	Fisheries Research and Monitoring Division, PIFSC
FRS	DAR Fisher-Reporting System data set
JIMAR	Joint Institute for Marine and Atmospheric Research
KIR	Kahoolawe Island Reserve
LHP	Life History Program, FRMD, PIFSC
MHI	Main Hawaiian Islands
NMFS	National Marine Fisheries Service, a.k.a. NOAA Fisheries
NOAA	National Oceanic and Atmospheric Administration
NWHI	Northwest Hawaiian Islands
PIFSC	Pacific Islands Fisheries Science Center, NMFS, NOAA
PIRO	Pacific Islands Regional Office, NMFS, NOAA
SAP	Stock Assessment Program, FRMD, PIFSC
TAC	Total Allowable Catch
WRPFMC	Western Regional Pacific Fishery Management Council

Table 12. List of species in this report by common name, scientific name, type of fishery classification, and Hawaii Department of Aquatic Resources (DAR) species code

Name	Scientific name	Type	Species code
OPAKAPAKA	*Pristipomoides filamentosus*	Deep 7 bottomfish	19
ONAGA	*Etelis coruscans*	Deep 7 bottomfish	22
KALEKALE	*Pristipomoides sieboldii*	Deep 7 bottomfish	17
EHU	*Etelis carbunculus*	Deep 7 bottomfish	21, 36*
GINDAI	*Pristipomoides zonatus*	Deep 7 bottomfish	97
LEHI	*Aphareus rutilans*	Deep 7 bottomfish	58
HAPUUPUU	*Hyporthodus quernus*	Deep 7 bottomfish	15
UKU	*Aprion virescens*	Non-deep 7 bottomfish	20
KAHALA	*Seriola dumerili*	Non-deep 7 bottomfish	16
TAAPE	*Lutjanus kasmira*	Non-deep 7 bottomfish	114
PAPIO, ULUA (MISC.)	*Carangidae*	Coral Reef Ecosystem	23
TUNA (MISC.)	*Scombridae*	Pelagic	1
AKU	*Katsuwonus pelamis*	Pelagic	2
YELLOWFIN TUNA	*Thunnus albacares*	Pelagic	3
TOMBO	*Thunnus alalunga*	Pelagic	4
BLUEFIN TUNA	*Thunnus thynnus*	Pelagic	5
BIGEYE TUNA	*Thunnus obesus*	Pelagic	6
KAWAKAWA	*Euthynnus affinis, E. yaito*	Pelagic	7
BILLFISH (MISC.)	*Istiophoridae and Xiphiidae*	Pelagic	8
STRIPED MARLIN	*Tetrapterus audax*	Pelagic	9
BLUE MARLIN	*Makaira mazara, M. nigricans*	Pelagic	10
SWORDFISH	*Xiphias gladius*	Pelagic	11
SAILFISH	*Istiophorus platypterus*	Pelagic	12
MAHIMAHI	*Coryphaena hippurus*	Pelagic	13
ONO	*Acanthocybium solandri*	Pelagic	14
HAULIULI	*Gempylus serpens*	Pelagic	39
OPAH	*Lampris guttatus*	Pelagic	106
SHORTBILL SPEARFISH	*Tetrapterus angustirostris*	Pelagic	107
BLACK MARLIN	*Makaira indica*	Pelagic	108
MONCHONG	*Taractichthys (Taractes) steindachneri*	Pelagic	118
MAKO	*Isurus oxyrinchus*	Pelagic	320
THRESHER SHARK	*Alopias vulpinus*	Pelagic	321
BLUE SHARK	*Euselachii*	Pelagic	323
OCEANIC WHITETIP SHARK	*Euselachii*	Pelagic	324

* Code 36 only used in early 1950s and discontinued in 1989.

APPENDIX C: Participant List and Workshop Agendas

Participant List **Present at workshop#**

FIRST NAME	LAST NAME	AFFILIATION	1	2	3	4	5
Nathan	Abe	Commercial fisher	X	X	X	X	X
Jerry	Ault	University of Miami	-	-	X	X	-
Chris	Boggs	Pacific Islands Fisheries Science Center	X	X	X	X	X
Jon	Brodziak	Pacific Islands Fisheries Science Center	X	-	-	-	-
Stefanie	Dukes	Pacific Islands Fisheries Science Center	-	-	X	X	X
Matthew	Dunlap	Pacific Islands Regional Office	X	X	-	X	X
Edwin	Ebisui II	Western Pacific Regional Fishery Management Council	-	-	-	-	-
Edwin	Ebisui III	Commercial fisher	X	-	X	X	X
Eric	Fletcher	Pacific Islands Fisheries Science Center, Joint Institute for Marine and Atmospheric Research	X	X	-	X	X
Erik	Franklin	University of Hawaii	-	-	-	-	-
Kimberlee	Harding	Hawaii Department of Land and Natural Resources	-	X	X	-	X
Walter	Ikehara	Pacific Islands Regional Office	X	X	X	-	X
David	Itano	Western Pacific Regional Fishery Management Council's Scientific and Statistical Committee	X	X	-	-	-
Ariel	Jacobs	Pacific Islands Regional Office	-	-	-	-	-
Kurt	Kawamoto	Pacific Islands Fisheries Science Center	X	X	X	X	X
Don	Kobayashi	Pacific Islands Fisheries Science Center, Western Pacific Regional Fishery Management Council's Scientific and Statistical Committee	X	X	X	X	-
Reginald	Kokubun	Hawaii Department of Land and Natural Resources	X	X	X	X	X
Brian	Langseth	Pacific Islands Fisheries Science Center	-	-	X	X	X
Kimberly	Lowe	Pacific Islands Fisheries Science Center	X	X	-	-	-
Beth	Lumsden	Pacific Islands Fisheries Science Center	X	X	X	X	X
Mark	Mitsuyasu	Western Pacific Regional Fishery Management Council staff	X	-	X	-	-
Alton	Miyasaka	Hawaii Department of Land and Natural Resources	-	-	-	-	-
Roy	Morioka	Waialua Boat Club, Commercial fisher	X	X	X	X	X
Layne	Nakagawa	Commercial fisher	X	-	X	X	X
Basil	Oshiro	Commercial fisher, Maui Cooperative Fishing Association	X	-	-	-	-
Zack	Oyafuso	University of Hawaii	-	-	-	-	X
Marlowe	Sabater	Western Pacific Regional Fishery Management Council staff	X	X	X	X	X
Michael	Seki	Pacific Islands Fisheries Science Center	-	X	X	X	X
Steve	Smith	University of Miami	-	-	X	X	-
Clay	Tam	Pacific Islands Fisheries Group	X	X	X	X	X
Edwin	Watamura	Waialua Boat Club, Commercial fisher	X	-	X	X	X
Leonard	Yamada	Aiea Boat Club, Commercial fisher	X	X	X	X	X
Annie	Yau	Pacific Islands Fisheries Science Center	X	X	X	X	X

Workshop 1 Agenda

Bottomfish Commercial Fishery Data Workshop

Workshop 1

September 2, 2015

NOAA Pier 38 conference room

8 AM – 4 PM

1. Welcome and Introductions
2. Background and Terms of Reference
3. DAR data issues
    a. Catch data
    b. Weight data
    c. CPUE

Workshop 2 Agenda

Bottomfish Commercial Fishery Data Workshop

On the Hawaii Division of Aquatic Resources (DAR),
Commercial Fisher Reporting System (FRS) Data

Workshop 2

November 4–5, 2015
NOAA Pier 38 conference room
8 AM – 4 PM each day

The focus of this workshop will be on data issues surrounding calculation of nominal catch-per-unit-effort (CPUE).

There will be no-host lunch breaks, and at least 2 additional breaks scattered throughout each day.

Timing is approximated, and the meeting may adjourn earlier if all agenda items are completed.

DAY 1

1. Welcome and Introductions
2. Background and Terms of Reference (reminder)
3. DAR data issues
    a. Catch data – recap of action items from Workshop 1, future steps
    b. Weight data – recap of action items from Workshop 1, future steps

LUNCH

    c. CPUE
        i. Discuss issues
        ii. Determine action items and assign to individuals

DAY 2

    c. CPUE continued if needed…
        i. Discuss issues
        ii. Determine action items and assign to individuals

LUNCH

    c. CPUE continued if needed…
        iii. Discuss issues
        iv. Determine action items and assign to individuals

Workshop 3 Agenda

Bottomfish Commercial Fishery Data Workshop
On the Hawaii Division of Aquatic Resources (DAR),
Commercial Fisher Reporting System (FRS) Data

Workshop 3
April 26–27, 2015
NOAA Pier 38 conference room
8 AM – 4 PM each day

The focus of this workshop will be to:
1) Wrap up and finalize any remaining issues with calculation of catch data, and
2) Discuss and prioritize exploratory analyses in addressing issues surrounding calculation of nominal catch-per-unit-effort (CPUE) stemming from discussion at Workshop 2.

There will be no-host lunch breaks, and at least 2 additional breaks scattered throughout each day.

Timing is approximated, and the meeting may adjourn earlier if all agenda items are completed.

DAY 1

1. Welcome and Introductions
2. Background and Terms of Reference (reminder)
3. Introduce draft report document
4. DAR data issues
    a. Catch data
        i. Recap of action items from Workshop 2
        ii. FINALIZE method for calculating catch

LUNCH

    b. Weight data
        i. Recap of action items from Workshop 2, future steps
    c. CPUE data
        i. Determine analyses required to resolve CPUE steps
        ii. Assign action items to individuals
        iii. Presentation from guest scientists, Jerry Ault and Steve Smith of University of Miami

DAY 2

    c. CPUE data continued as needed…
        i. Determine analyses required to resolve CPUE steps
        ii. Assign action items to individuals

LUNCH

    c. CPUE data continued as needed…
        i. Determine analyses required to resolve CPUE steps
        ii. Assign action items to individuals

Workshop 4 Agenda

Bottomfish Commercial Fishery Data Workshop
On the Hawaii Division of Aquatic Resources (DAR),
Commercial Fisher Reporting System (FRS) and Dealer Reporting System (DRS) Data

Workshop 4
August 1–5, 2016
14th floor conference room, Western Pacific Regional Fishery Management Council
(17th floor conference room on Wednesday)
8:30 AM – 5 PM

The focus of this workshop will be on CPUE steps 1-3:

1) Selecting bottomfishing data records,
2) Selecting records that comply with a single unit of effort, and
3) Selecting data most representative of CPUE trends over time.

A small group of analysts identified at the 3rd workshop has worked intersessionally on CPUE steps 1-3 using input provided by everyone at past 3 workshops. At this 4th workshop, they will present results for group discussion and decision-making.

There will be no-host lunch breaks, and at least 2 additional breaks scattered throughout each day.
Timing is approximated, and the meeting may adjourn earlier if all agenda items are completed.

DAY 1, MONDAY

1. Welcome, Introductions, Logistics
2. Background and Terms of Reference (reminder)
3. Comments on draft report document?
4. CPUE Step 1
    a. Presentation
    b. Discussion
    c. Any requested additional analyses?

LUNCH

5. CPUE Step 2
    a. Presentation
    b. Discussion
    c. Any requested additional analyses?

DAY 2, TUESDAY

(Workshop does not convene. Analysts will work on new analysis requests.)

(Ben Richards will provide a short presentation and Q&A session on fishery-independent survey design and analysis for interested parties)

DAY 3, WEDNESDAY (17th floor conference room)

4. CPUE Step 1
    a. Presentation of new analyses
    b. Discussion
    c. Final decisions on Step 1

5. CPUE Step 2
    a. Presentation of new analyses
    b. Discussion
    c. Final decisions on Step 2

LUNCH

6. CPUE Step 3
    d. Presentation
    e. Discussion
    f. Any requested additional analyses?

DAY 4, THURSDAY

(Workshop does not convene. Analysts will work on new analysis requests.)
(If needed, Ben Richards can continue to answer questions on fishery-independent survey)

DAY 5, FRIDAY

6. CPUE Step 3
    a. Presentation of new analyses
    b. Discussion
    c. Final decisions on Step 3

LUNCH

7. Extra time as needed to wrap up lingering issues with CPUE steps 1-3

Workshop 5 Agenda

Bottomfish Commercial Fishery Data Workshop
On the Hawaii Division of Aquatic Resources (DAR),
Commercial Fisher Reporting System (FRS) and Dealer Reporting System (DRS) Data

Workshop 5
November 14-18– 2016 (Monday, Wednesday, Friday)
NOAA Pier 38 conference room
8:00 AM – 3:45 PM each day

The focus of this workshop will be on weight. Information on fish weight over time is used to infer changes in the stock over time.
Weight data needs to represent what is happening to the stock in the water.
The weight data TREND (increasing vs decreasing) and SCALE (exact values in pounds) are important.

GOAL: Agree on filtering procedures to produce accurate, reliable, and representative weight data for stock assessment use. We do not need to choose a single set of weight data. Instead, we will choose all the weight data that can be used for stock assessments, and then only a specific set will be used depending on the stock assessment model chosen later.

A small group of analysts has worked intersessionally on weight data analyses using input provided by everyone at past workshops. At this 5th and final workshop, they will present results for group discussion and decision-making.

There will be no-host lunch breaks, and at least 2 additional breaks scattered throughout each day.

Timing is approximated, and the meeting may adjourn earlier if all agenda items are completed.

DAY 1, MONDAY

1. Welcome, Introductions, Logistics
2. Background and Terms of Reference (reminder)
3. Draft report document update
4. Go over handout to guide decisions on weight data
5. Fisher-reporting system (FRS) weight data
    a. Presentation, discussion, decisions
    b. Any requested additional analyses?

LUNCH

6. Fisher-reporting system (FRS) weight data
    a. Presentation, discussion, decisions
    b. Any requested additional analyses?

DAY 2, TUESDAY

(Workshop does not convene. Analysts will work on new analysis requests.)
(8-10 AM: OPTIONAL: Ben Richards available to provide a short presentation and Q&A session on fishery-independent survey design and analysis for interested parties; he will also meet with Clay Tam, Reggie Kokubun, Kimberly Harding, and Annie Yau to discuss best future practices for reporting bottomfish survey fish and effort.)

DAY 3, WEDNESDAY

7. Fisher-reporting system (FRS) weight data
    a. Presentation, discussion
    b. Final decisions completed for FRS data

LUNCH

8. Dealer-reporting system (DRS) weight data
    g. Presentation, discussion, decisions
    h. Any requested additional analyses?

DAY 4, THURSDAY

(Workshop does not convene. Analysts will work on new analysis requests.)

DAY 5, FRIDAY

9. Dealer-reporting system (DRS) weight data
    a. Presentation, discussion, decisions

LUNCH

10. ALL FINAL DECISIONS made for weight data
11. Closeout of Bottomfish Commercial Fishery Data workshops

APPENDIX D.1: R Code to Implement Catch Data Resolutions for Stock Assessments

```
########################################
#Purpose: Calculated catch by deep7 species by fishing year
#Use: 2018 Deep 7 MHI bottomfish assessment
#By: Brian Langseth, PIFSC
#Created: 12/19/2016
#Last updated: 12/20/2016
#Rversion 3.2.4

#Input: Files provided by Paul Tao on June 28, 2016
#       updating license names and number from ORACLE DAR data back to 1948.
#      Also fixed some duplicate license numbers.

#Output: Catch of deep7 bottomfish species for MHI
#       grids only by fishing year and species
########################################

rm(list=ls())

# --Combine input data sets into single data set--------------------------
library(foreign)

early=read.dbf("D:\\FilePath\\HFY48_93E.dbf",as.is=T)
late=read.dbf("D:\\FilePath\\hf94c15F.dbf",as.is=T)

#Add fields added by Paul in late file to early file.
#Initially not provided in early, so original entries for all were NA
early[,c("O_LICENSE","O_LNAME","O_FNAME","O_VESSEL","O_USCG","O_HA_NO","
CHNG")]=NA
fulldata=rbind(early,late)
dim(fulldata)

#Add lbs field as maximum of LBS and LBS_SOLD
fulldata$lbs=pmax(fulldata$LBS,fulldata$LBS_SOLD,na.rm=T)

# --Use only records of Deep 7 species ----------------------------------
##Read in species names from Oracle download
fish_spec=read.csv("D:\\FilePath\\fisher_species_5_6_16.csv",header=T)

#Use common names of deep7 species since Gindai has two scientific names
com_names=c("HAPUUPUU","KALEKALE","OPAKAPAKA","EHU","ONAGA","ehu","LEH
I","GINDAI")
```

```
##Species code 36 is also ehu according to Appendix II.4g in Moffitt et al. 2011
# "Bottomfish CPUE standardization workshop proceedings August 4-6, 2008",
# but was discontinued in 1989.
id=c(15,17,19,21,22,36,58,97)
fish_spec[fish_spec$SPECIES_PK%in%id,]

####----#####
deep7=fulldata[fulldata$SPECIES%in%id,]
dim(deep7)
####----#####

#Assign ehu species code 36 to ehu code 21.
deep7[deep7$SPECIES==36,"SPECIES"]=21

# --Use only records from the MHI-----------------------------------
#Read in the mhi_areas excel file from Reggie that lists which are
#valid entries and which are not, and which are ponds
#Only keep valid entries
areasReggie=read.csv("D:\\FilePath\\BF_Area_Grid_Reggie.csv",header=T)
valid=areasReggie[which(areasReggie$Valid.==""),]$area

#Remove non-valid subareas (A and B) from area 16123 as well as
#records from 16123 without a subarea specified
####----#####
mhidata=deep7[deep7$AREA%in%valid,] #valid areas
mhidata=mhidata[!mhidata$SUBAREA%in%c("A","B"),] #remove known invalid subareas
(16123A and 16123B)
mhidata=mhidata[!(mhidata$AREA==16123 & is.na(mhidata$SUBAREA)),] #remove the
16123 records without a subarea distinction
####----#####

# --Partition invalid areas to known mhi ares----------------------------
#Based on proportion of species/year in mhi area compared to nwhi areas
#Invalid areas include 0's, ponds, and non valid MHI areas
invalid=areasReggie[which(areasReggie$Valid.!=""),]$area
notvalid=deep7[deep7$AREA%in%invalid,]

#Need NWHI areas as well
#NWHI areas are those not in mhi_areas.csv
####----#####
nwhidata=deep7[!deep7$AREA%in%areasReggie$area,]
####----#####

#Calculate proportion of fish by species and year in MHI vs NWHI
```

```
agg_mhi=aggregate(mhidata$lbs,by=list("FYEAR"=mhidata$FYEAR,"SPECIES"=mhidata$SP
ECIES),FUN=sum)
mhi_table=matrix(agg_mhi$x,length(1948:2015),length(id[-6]),dimnames=list(1948:2015,id[-
6]))

agg_nwhi=aggregate(nwhidata$lbs,by=list("FYEAR"=nwhidata$FYEAR,"SPECIES"=nwhidata
$SPECIES),FUN=sum)
comb_agg=merge(agg_mhi,agg_nwhi,by=c("FYEAR","SPECIES"),all.x=T)
comb_agg[is.na(comb_agg$x.y),"x.y"]=0
comb_agg$percMHI=comb_agg$x.x/(comb_agg$x.x+comb_agg$x.y)

agg_inval=aggregate(notvalid$lbs,by=list("FYEAR"=notvalid$FYEAR,"SPECIES"=notvalid$S
PECIES),FUN=sum)
comb_all=merge(comb_agg,agg_inval,by=c("FYEAR","SPECIES"),all=T)
comb_all$add_catch=comb_all$percMHI*comb_all$x

invalid_ADD=matrix(comb_all$add_catch,length(1948:2015),length(id[-
6]),dimnames=list(1948:2015,id[-6]))
invalid_ADD[is.na(invalid_ADD)]=0

#Final reported catch table
final_reported_table=round(invalid_ADD+mhi_table,2)

# --Output material--------------------------------------
#Final data set
#write.csv(mhidata,"D:\\FilePath\\Finalized_catch_dataset.csv",row.names=F)

#Final reported table
#write.csv(final_reported_table,"D:\\FilePath\\Reported_catch_table.csv",row.names=T)
```

APPENDIX D.2: R Code to Implement Fish Weight Data Resolutions for Stock Assessments

```
########################################
#Purpose: Prepare and calculate weight data
#        from FRS based on decisions made at the weight data workshop in Nov, 2016
#Use: 2018 Deep 7 MHI bottomfish assessment
#By: Brian Langseth, PIFSC
#Created: 11/21/2016
#Last updated: 8/3/2016
#Rversion 3.2.4

#Input: Original FRS files provided by Paul Tao on June 28, 2016.
#
#Output: Filtered data set based on workshop
#      decisions, and calculated weight by year
#      with std. error for opakapaka/onaga using
#      all records as decided on at weight
#      workshop in Nov. 2016.
########################################

rm(list=ls())
library(plyr)
library(foreign)

# --Read in initial data sets from Paul Tao and combine into single data set--------------------------
early=read.dbf("D:\\FilePath\\HFY48_93E.dbf",as.is=T)
late=read.dbf("D:\\FilePath\\hf94c15F.dbf",as.is=T)

#Add fields added by Paul in late file to early file.
#Initially not provided in early, so original entries for all were NA
early[,c("O_LICENSE","O_LNAME","O_FNAME","O_VESSEL","O_USCG","O_HA_NO","
CHNG")]=NA
#Combine data sets
fulldata=rbind(early,late)
dim(fulldata)

# --Filter out records--------------------
#Reduce to only deep7 species
id=c(15,17,19,21,22,36,58,97)
com_names=c("HAPUUPUU","KALEKALE","OPAKAPAKA","EHU","ONAGA","ehu","LEH
I","GINDAI")
frs1=fulldata[fulldata$SPECIES%in%id,]
dim(frs1)
#assign species code 36 (4065 records) to be 21 (both are ehu)
frs1[frs1$SPECIES==36,"SPECIES"]=21
#Read in the mhi_areas excel file from Reggie that lists which are
```

```
#valid entries and which are not, and which are ponds
#Only keep valid entries
areasReggie=read.csv("D:\\FilePath\\BF_Area_Grid_Reggie.csv",header=T)
valid=areasReggie[which(areasReggie$Valid.==""),]$area
frs2=frs1[frs1$AREA%in%valid,]
dim(frs2)
#Filter out A and B subareas for area 16123, which is the only region with
#subareas. Remove non-subareas in 16123. Only 16123-C is valid for 16123.
table(frs2[!is.na(frs2$SUBAREA),"AREA"])
frs2=frs2[!frs2$SUBAREA%in%c("A","B"),] #remove A and B subareas
frs2=frs2[!(frs2$AREA==16123 & is.na(frs2$SUBAREA)),] #remove the 16123 records without
a subarea distinction

#Remove partial fishing year 2016 records.
frs3=frs2[frs2$FYEAR<2016,]
dim(frs3)

#Keep only gear 3 records.
frs4=frs3[frs3$GEAR==3,]
dim(frs4)

#Filter out records with either 0 lbs or 0 number.
frs5=frs4[frs4$LBS>0 & frs4$CAUGHT>0,] #Number sold is too irregularly used
dim(frs5)

#Remove records with extreme average weights as based on max and min
#weights from the DAR species list
specTAB_f=read.csv("D:\\FilePath\\fisher_species_5_6_16.csv",header=T)
specTAB_d=read.csv("D:\\ FilePath \\dealer_species_5_6_16.csv",header=T)
deep7_f=specTAB_f[specTAB_f$SPECIES_PK%in%id,c("SPECIES_PK","MIN_WT_CAUGH
T","MAX_WT_CAUGHT")]
deep7_d=specTAB_d[specTAB_d$SPECIES_PK%in%id,c("SPECIES_PK","MIN_WT_CAUG
HT","MAX_WT_CAUGHT")]
#Merge min(frs) and max(drs) weights to main data set
temp_min=merge(frs5,deep7_f[,c(1,2)],by.x="SPECIES",by.y="SPECIES_PK")
frs6=merge(temp_min,deep7_d[,c(1,3)],by.x="SPECIES",by.y="SPECIES_PK")
#Calculate average weight
frs6$avgw=frs6$LBS/frs6$CAUGHT
#Filter out records with avg weight > max and avg weight < min
out_f=c(which(frs6$avgw<frs6$MIN_WT_CAUGHT),which(frs6$avgw>frs6$MAX_WT_CAU
GHT))
length(out_f)
frs7=frs6[-out_f,]
dim(frs7)

# --Output final full data set--------------------
wdata=frs7 #final weight data set
```

```
#write.csv(wdata,"D:\\FilePath\\Finalized_weight_dataset.csv",row.names=F)

# --Weight calculations--------------------
#Calculate average weight of opakapaka and onaga together using all records.

#Need a weighted standard deviation in order to calculate
#standard error. Use my own function based on the frequency-weights
#(repeat-weights) way of calculating weighted variance, and is unbiased
wgt.var=function(data,weights){
  wmean=sum(data*weights)/sum(weights) #weighted mean
  wvar=sum(weights*(data-wmean)^2)/(sum(weights)-1)
  return(wvar)
}

wdata$avgw_kg=wdata$avgw/2.20462 #convert lbs to kg (for SS)

#For paka and onaga
wavg = ddply(wdata[wdata$SPECIES%in%c(19,22),], .(FYEAR), summarize,
mean_w=weighted.mean(avgw,CAUGHT), sd_w=sqrt(wgt.var(avgw,CAUGHT)),
n=length(avgw),fish=sum(CAUGHT),se=sqrt(wgt.var(avgw,CAUGHT)/sum(CAUGHT)))
#Remove first data point (1948) due to being a half year and first
#part of data set and for looking very different than other years
wavg2=wavg[-1,]

wavgkg = ddply(wdata[wdata$SPECIES%in%c(19,22),], .(FYEAR), summarize,
mean_w_kg=weighted.mean(avgw_kg,CAUGHT),
sd_w_kg=sqrt(wgt.var(avgw_kg,CAUGHT)),
n=length(avgw_kg),fish=sum(CAUGHT),se_kg=sqrt(wgt.var(avgw_kg,CAUGHT)/sum(CAUG
HT)))
wavgkg2=wavgkg[-1,]

#For paka only
wavg_paka = ddply(wdata[wdata$SPECIES%in%c(19),], .(FYEAR), summarize,
mean_w=weighted.mean(avgw,CAUGHT), sd_w=sqrt(wgt.var(avgw,CAUGHT)),
n=length(avgw),fish=sum(CAUGHT),se=sqrt(wgt.var(avgw,CAUGHT)/sum(CAUGHT)))
wavg_paka_nonexp = ddply(wdata[wdata$SPECIES%in%c(19),], .(FYEAR), summarize,
mean_w_nonexp=mean(avgw), sd_w_nonexp=sqrt(var(avgw)),
n=length(avgw),se_nonexp=sqrt(var(avgw)/length(avgw)))
#Remove first data point (1948) due to being a half year and first
#part of data set and for looking very different than other years
wavg2paka=wavg_paka[-1,]
wavg2paka_nonexp=wavg_paka_nonexp[-1,]
```

```
wavgkg_paka = ddply(wdata[wdata$SPECIES%in%c(19),], .(FYEAR), summarize,
mean_w_kg=weighted.mean(avgw_kg,CAUGHT),
sd_w_kg=sqrt(wgt.var(avgw_kg,CAUGHT)),
n=length(avgw_kg),fish=sum(CAUGHT),se_kg=sqrt(wgt.var(avgw_kg,CAUGHT)/sum(CAUG
HT)),se_kg_n=sqrt(wgt.var(avgw_kg,CAUGHT)/length(avgw_kg)))
wavgkg2_paka=wavgkg_paka[-1,]
wavgkg_paka_nonexp = ddply(wdata[wdata$SPECIES%in%c(19),], .(FYEAR), summarize,
mean_w_kg_nonexp=mean(avgw_kg), sd_w_kg_nonexp=sqrt(var(avgw_kg)),
n=length(avgw_kg),se_kg_n_nonexp=sqrt(var(avgw_kg)/length(avgw_kg)))
wavgkg2_paka_nonexp=wavgkg_paka_nonexp[-1,]

# --Output final yearly averaged data set--------------------
#For paka and onaga
#write.csv(cbind(wavg2,"",wavgkg2),"D:\\FilePath\\Weight_timeseries.csv",row.names=F)
#For PAKA only
#write.csv(cbind(wavg2paka,"",wavgkg2_paka,"",wavg2paka_nonexp,"",wavgkg2_paka_nonex
p),"D:\\FilePath\\Weight_timeseries_PAKA.csv",row.names=F)
```

APPENDIX D.3: R Code to Implement CPUE Data Resolutions for Stock Assessments

```
########################################
#Purpose: Setup data set for CPUE analysis
#        for bottomfish assessment. This is a working document based on
#        decisions made in the data workshops.
#Use: 2018 Deep 7 MHI bottomfish assessment
#By: Brian Langseth, PIFSC
#Created: 8/8/2016
#Last updated: August 2017

#Rversion 3.2.4

#Input: Files provided by Paul Tao on June 28, 2016
#      updating license names and number from
#      DAR data back to 1948. Also fixed some
#      duplicate license numbers.

#Output: Best available data set for CPUE time series
#       I added fields: trip, lbs, deep7_lbs, deep7_perc,
#       n_ports, n_areas, uku_lbs, cum_exp,

########################################

rm(list=ls())

# --Combine input data sets into single data set-------------------------
library(foreign)
early=read.dbf("D:\\FilePath\\HFY48_93E.dbf",as.is=T)
late=read.dbf("D:\\FilePath\\hf94c15F.dbf",as.is=T)

#Add fields added by Paul in late file to early file.
#Initially not provided in early, so original entries for all were NA
#"O" refers to original entry
#"CHNG" refers to what was changed
early[,c("O_LICENSE","O_LNAME","O_FNAME","O_VESSEL","O_USCG","O_HA_NO","
CHNG")]=NA
#Combine data sets
fulldata=rbind(early,late)
dim(fulldata)

# --Use only records from the MHI-----------------------------
#Read in key table from last assessment for MHI areas and keep
#only these grids
mhi_area=read.csv("D:\\FilePath\\mhi_area.csv",header=T)
```

```
####----#####
mhidata=fulldata[fulldata$AREA%in%mhi_area$area,]
####----#####

dim(fulldata)-dim(mhidata) #removes 190,760 records from NWHI

# --Separate out deepsea handline gear as well as -------------------------
#other gears used on trips using deepsea handline gear
check_gear=table(mhidata$GEAR,useNA="always")

#Gear 3 records...
temp=mhidata[mhidata$GEAR==3,]
dim(temp)
#...add in records from other gears on "trips" that used gear 3

#Determine ‘single reporting day’ number for single reporting days that used gear 3;
# ‘trip’ used synonymously with ‘single reporting day’ in this code
addrec=unique(temp[,c("LICENSE","FISHED")])
#assign single reporting day number and merge it to data set
addrec$trip=c(1:nrow(addrec)) #assign trip number
temp2=merge(addrec,mhidata,by=c("LICENSE","FISHED")) #merge to data set
dim(temp2)
check_gear2=table(temp2$GEAR,useNA="always")

####----#####
#Records from gear 3 and other gears on trips with gear 3
mhidata2=temp2 #1011510 records
####----#####

#Remove partial fishing year 2016 records
####----#####
mhidata3=mhidata2[mhidata2$FYEAR<=2015,]
dim(mhidata3) #1006952 records
####----#####

# --Output final data sets -------------------------
#For group step 3 comparison
#Keep some personal information
temp_out=subset(mhidata3,select=-
c(USCG,HA_NO,IMP_USER,O_FNAME,O_LNAME,O_VESSEL,O_USCG,O_HA_NO,ORI
G_USER))
temp_out=temp_out[temp_out$GEAR==3,] #826,839 records
write.csv(temp_out,"D:\\FilePath\\ForGroupCpueAnalysis_gear3Step3.csv",row.names=F)

# --Read in basic reduced data set from 01_WORKSHOP_data processing--------------------------
datafull=read.csv("D:\\FilePath\\ForGroupCpueAnalysis_gear3Step3.csv",header=T)
dim(datafull)
```

```
# --FILTER data: keep duplicates, no invalid areas, no license number 0----------
a=duplicated(datafull)
sum(a) #663

#-Read in the mhi_areas excel file from Reggie that lists which are
#valid entries and which are not, and which are ponds
#Only keep valid entries
areasReggie=read.csv("D:\\FilePath\\BF_Area_Grid_Reggie.csv",header=T)
valid=areasReggie[which(areasReggie$Valid.==""),]$area
#Confirm subareas are valid for 16123 (which is a valid area)
table(datafull[!is.na(datafull$SUBAREA),]$AREA) #all subareas are for 16123
table(datafull$SUBAREA) #filter out A and B subareas as well as those without a subarea
####----#####
data=datafull[datafull$AREA%in%valid,] #valid areas
data=data[!data$SUBAREA%in%c("A","B"),] #remove known invalid subareas (16123A and
16123B)
data=data[!(data$AREA==16123 & is.na(data$SUBAREA)),] #remove the 16123 records
without a subarea distinction
####----#####

#-Remove all instances where license number is 0 because
#these are unlikely to be from the same trip
head(table(data$LICENSE))
table(data[data$LICENSE==0,]$FYEAR) #all in 1997 or before
####----#####
data=data[data$LICENSE>0,]
####----#####
dim(data) #800130 records

# --STEP01: select deep7 bottomfish records----------------------------

##-##
#-Determine the percentage of deep7 caught for each trip and add to data set
#Base this on the larger of lbs_kept and lbs_sold (see catch data for details)
id=c(15,17,19,21,22,36,58,97) #include both ehu codes
data$lbs=pmax(data$LBS,data$LBS_SOLD,na.rm=T)
alllbs=aggregate(data$lbs,by=list(data$trip),sum)
deep7lbs=aggregate(data[data$SPECIES%in%id,]$lbs,by=list(data[data$SPECIES%in%id,]$trip
),sum) #lbs
#merge records since many trips dont catch deep7
lbs_trip=merge(alllbs,deep7lbs,by="Group.1",all=T)
colnames(lbs_trip)=c("trip","all_lbs","deep7_lbs")
lbs_trip$deep7_perc = round(lbs_trip$deep7_lbs/lbs_trip$all_lbs,3)
#replace NAs with 0
lbs_trip[is.na(lbs_trip$deep7_perc),"deep7_perc"]=0
lbs_trip[is.na(lbs_trip$deep7_lbs),"deep7_lbs"]=0
dim(lbs_trip) #number of unique trips
```

```
####----#####
#Add percentage of deep7 fish caught per trip to database
data1=merge(data,lbs_trip[,c("trip","deep7_perc","all_lbs","deep7_lbs")],by=c("trip"))
####----#####

##-##
#-Remove some trips where no deep7 bottomfish were caught
#Of those trips that did not catch deep7, remove all records from
#trips that caught PMUS species, Uku, or Species=0 from Sept. 2002 back
#All records w/o deep7
data1old=data1[substr(data1$FISHED,1,7)<="2002-09",]
all0=data1old[data1old$deep7_perc==0,]
table(all0$SPECIES%in%id) #12 records where deep7 were recorded, just not as weight
dim(all0)
#Records from trips that caught PMUS
pmus.id=c(1:14,106,107,108,320,321,323,324,39,118) #39 is snake mackeral (Gempylidae) and
118 is monchong (Bramidae)
pmus_all0=all0[all0$SPECIES%in%pmus.id,] #records of PMUS from trips w/o deep7
pmus_all0_trips=unique(pmus_all0$trip) #trips with records of PMUS from trips w/o deep7
pmus_all0_records=all0[all0$trip%in%pmus_all0_trips, ]#all records from trips with PMUS w/o
deep7
dim(pmus_all0_records)
#Records from trips that caught Uku
uku_all0=all0[all0$SPECIES==20,] #records of Uku from trips w/o deep7
uku_all0_trips=unique(uku_all0$trip) #trips with records of Uku from trips w/o deep7
uku_all0_records=all0[all0$trip%in%uku_all0_trips,] #all records from trips with uku w/o deep7
dim(uku_all0_records)
#Records from trips that had Species=0
spec0_all0=all0[all0$SPECIES==0,] #records of Species=0 from trips w/o deep7
spec0_all0_trips=unique(spec0_all0$trip) #trips with records of Species=0 from trips w/o deep7
spec0_all0_records=all0[all0$trip%in%spec0_all0_trips,] #all records from trips with Species=0
w/o deep7
dim(spec0_all0_records)
#Total number of records removed from records w/o deep7
dim(all0[all0$trip%in%unique(c(pmus_all0_trips,uku_all0_trips,spec0_all0_trips)),])
####----#####
#Remove trips w/o deep7, and with records of PMUS species, uku, and species=0 <= Sept. 2002
data2=(data1[!(data1$deep7_perc==0 & data1$trip%in%pmus_all0_trips) &
        !(data1$deep7_perc==0 & data1$trip%in%uku_all0_trips) &
        !(data1$deep7_perc==0 & data1$trip%in%spec0_all0_trips),])
####----#####
```

```
##-##
#-Kona/South Point: Remove trips that caught any PMUS species while catching <50lbs of deep7 in FY1985 and before
BIarea = c(108,100,101,102,128,120,121,122 ) # DAR grids off Kona and South Point
BIdata = data2[data2$AREA%in%BIarea,]
#Which of these trips had deep7 catch less than 50 lbs and caught a PMUS in FY1985 and before
BIdata_pmus=BIdata[BIdata$SPECIES%in%pmus.id,]
BIdata_pmus50=BIdata_pmus[BIdata_pmus$deep7_lbs<50,]
BIdata_pmus50_1985=BIdata_pmus50[BIdata_pmus50$FYEAR<=1985,]
####----#####
data3=data2[!data2$trip%in%unique(BIdata_pmus50_1985$trip),]
dim(data3)
####----#####

##-##
#-Remove records (trips) that occurred when the fishery was closed
#Date fishery closed: Apr 16,2008; Jul 6,2009; Apr 20,2010; Mar 12,2011
#Theres a lot, originally many of these were removed when I included
#trips after Sept. 2002 to be removed that didnt catch any deep7 (but
#did catch PMUS, Uku, and species=0) - most (by weight) of these are uku
data3$FISHED=as.Date(data3$FISHED)
####----#####
data4=data3[!(data3$FISHED>"2008-04-16" & data3$FISHED<="2008-08-31") &
            !(data3$FISHED>"2009-07-06" & data3$FISHED<="2009-08-31") &
            !(data3$FISHED>"2010-04-20" & data3$FISHED<="2010-08-31") &
            !(data3$FISHED>"2011-03-12" & data3$FISHED<="2011-08-31"),]
####----#####

# --STEP02: select records that comply to a unit of effort---------------------

#Single reproting day (CML-FISHED) combo up through Sept. 2002
#Hours fished after Sept. 2002

##-##
#-Multiday trips: distance cutoff
#Read in distance values
distances=read.csv("D:\\FilePath\\comport_area_distance_table.csv",header=T,col.names=c("com_port","AREA","SUBAREA","distance"))
ports=read.csv("D:\\FilePath\\port_comport_table.csv",header=T,col.names=c("PORT_LAND","com_port"))
port_changes=matrix(c(101,182,201,291,427,411),3,2,byrow=T)
colnames(port_changes)=c("PORT_LAND","com_port")
ports2=rbind(port_changes,ports)
####----#####
#Add common port to data set.
data4b=merge(data4[substr(data4$FISHED,1,7)<="2002-09",],ports2,by="PORT_LAND",all.x=T)
```

```
table(data4b$com_port,useNA="always") #5809 records that dont have a corresponding
com_port
####----#####
#Remove distance combination for subareas of area 16123 that aren't in the data
#which are all subareas other than subarea "C"
head(data4b[data4b$AREA==16123,])
distances=distances[!(distances$AREA==16123 & distances$SUBAREA!="C"),]

##Assign number of known common ports visited and number of areas visited on a trip
#How many areas per trip are there?
multiarea=rowSums(table(data4b$trip,data4b$AREA)>0)
table(multiarea)
#How many common ports per trip are there?
multiport=rowSums(table(data4b$trip,data4b$com_port)>0)
table(multiport,useNA="always") #161 trips visit 2 common ports (0 are because a common port
isn't available for PORT_LAND)
ports_v=data.frame("trip"=names(multiport),"n_com_ports"=multiport)
areas_v=data.frame("trip"=names(multiarea),"n_areas"=multiarea)
ports_areas=merge(ports_v,areas_v,by="trip",sort=FALSE)
####----#####
#number of known common ports and areas per trip
data4c=merge(data4b,ports_areas,by="trip")
####----#####
table(data4c$n_com_ports,data4c$n_areas)
table(unique(data4c[,c("trip","n_com_ports","n_areas")])$n_com_ports,unique(data4c[,c("trip","
n_com_ports","n_areas")])$n_areas)

##Assign distances based on combination of ports-areas
#FOR TRIPS WITH 1 AREA
data4c_1area=merge(data4c[data4c$n_areas==1,],distances[,-
3],by=c("com_port","AREA"),all.x=T)
#Those with 1 area and 1 common port. Distances are valid.
data4c_1area_1=data4c_1area[data4c_1area$n_com_ports==1,]
#
#Those with 1 area and 2 common ports.
#Need to separate trips when applicable and use a single distance
data4c_1area_2=data4c_1area[data4c_1area$n_com_ports>1,]
#Separate trips when applicable:
#For ports on same island or with general port code on same island, keep together as 1 trip.
#For ports on distant islands, separate. For ports with very distant
#island, keep together because likely a database error (define as greater than 3 islands away,
#so Kauai-Maui, or Oahu-Big Island).
u_areas=tapply(data4c_1area_2$com_port,data4c_1area_2$trip,unique)
u_areas2=matrix(as.numeric(unlist(u_areas,use.names=F)),length(u_areas),2,byrow=T,dimname
s=list(rownames(u_areas)))
sametrips=rownames(u_areas2)[which(substr(u_areas2[,1],1,1)==substr(u_areas2[,2],1,1))]
difftrips=rownames(u_areas2)[which(substr(u_areas2[,1],1,1)!=substr(u_areas2[,2],1,1))]
```

```
#Of the difftrips, which are very far apart (Big Island - Oahu or Maui - Kauai)
#and thus assumed to be errors. Want to keep together
far_difftrips=rownames(u_areas2)[which(abs(u_areas2[,1]-u_areas2[,2])>300)]
difftrips2=difftrips[!difftrips%in%far_difftrips]
#Take the trips that are to be separate, reclassify trip and add back to database
tosep=data4c_1area_2[data4c_1area_2$trip%in%difftrips2,]
tosep$trip2=paste0(tosep$trip,".",floor(tosep$com_port/100)) #to add a decimal that indicates
the first digit of the com_port (needed to distinguish ports 499 and 501, which weren't when
rounding)
data4c_1area_2sep=merge(data4c_1area_2,unique(tosep[,c("trip","com_port","trip2")]),by=c("tri
p","com_port"),all.x=T)
data4c_1area_2sep[is.na(data4c_1area_2sep$trip2),"trip2"]=data4c_1area_2sep[is.na(data4c_1ar
ea_2sep$trip2),"trip"]
#Reclassify distance.
#For each trip2 the most frequent distance is used, which can include NAs.
#Create own function to extract the modal distance. For instances where the number of distances
#is the same (e.g. two record trip that goes to two com_ports, this function defaults to the shorter
#distance)
modal=function(x){
  aa=as.numeric(names(which.max(table(x,useNA="always"))))
  return(aa)}
temp=aggregate(data4c_1area_2sep$distance,by=list(data4c_1area_2sep$trip2),modal)
colnames(temp)=c("trip2","distance2")
data4c_1area_2FINAL=merge(data4c_1area_2sep,temp,by="trip2")


#FOR TRIPS WITH 2+ AREA
data4c_manyareas=merge(data4c[data4c$n_areas>1,],distances[,-
3],by=c("com_port","AREA"),all.x=T)
#Those with many areas and 1 common port
data4c_manyareas_1=data4c_manyareas[data4c_manyareas$n_com_ports==1,]
head(data4c_manyareas_1[order(data4c_manyareas_1$trip),],20)
#Reclassify distance.
#Take greatest distance among areas as the trip distance
temp=aggregate(data4c_manyareas_1$distance,by=list(data4c_manyareas_1$trip),max)
colnames(temp)=c("trip","distance2")
data4c_manyareas_1FINAL=merge(data4c_manyareas_1,temp,by="trip")

#
#Those with many areas and 2 common ports
#Need to separate trips when applicable and use a single distance
data4c_manyareas_2=data4c_manyareas[data4c_manyareas$n_com_ports>1,]
head(data4c_manyareas_2[order(data4c_manyareas_2$trip),],20)
#Separate by trip. Nearly all have distinct areas for distinct ports
#and vice versa. Hence use trip-com_port as unique trip identifier for all
#Take the trips that are to be separate, reclassify trip and add back to database
```

```
data4c_manyareas_2$trip2=paste0(data4c_manyareas_2$trip,".",data4c_manyareas_2$com_port
) #use entire com_port to separate out trips
#Reclassify distance.
#Take max distance among areas as the trip distance
temp=aggregate(data4c_manyareas_2$distance,by=list(data4c_manyareas_2$trip2),max)
colnames(temp)=c("trip2","distance2")
data4c_manyareas_2FINAL=merge(data4c_manyareas_2,temp,by="trip2")

#FOR TRIPS WITH 0 COM_PORT
#No need to add distances since com_port is unavailable
data4c_0port=data4c[data4c$n_com_ports==0,]

##Combine all trips together into a single database
#This includes the records without com_ports
#Ensure all sub data sets have all names included
data4c_1area_1$trip2=data4c_1area_1$trip #Add trip2 to 1port-1area data set
data4c_1area_1$distance2=data4c_1area_1$distance #Add distance2 to 1port-1area data set
data4c_manyareas_1FINAL$trip2=data4c_manyareas_1FINAL$trip #Add trip2 to 1port-
manyarea data set
data4c_0port$distance=data4c_0port$distance2=NA #Add NA distance and distance2 to 0port
data set
data4c_0port$trip2=data4c_0port$trip #Add trip2 to 0port data set
##Single database
data4c_corrected=rbind(data4c_1area_1,data4c_1area_2FINAL,data4c_manyareas_1FINAL,dat
a4c_manyareas_2FINAL,data4c_0port)
#
#The 10 records in 1 area-2 ports trips with com_port
#not in database
table(data4c_corrected[is.na(data4c_corrected$distance2),]$n_com_ports)
#So NA's in distance are either due to a PORT_LAND that
#doesnt have a corresponding com_port, or a com_port/Area
#combination not in the database

##Now determine distance cutoff
#Corrected distances BY TRIP2 in 10yr increments
data4c_corrected_trip2=unique(data4c_corrected[,c("trip2","FYEAR","distance2")])
#Use 30nm for cutoff after 1960. Its inclusive of more data, and
#minimizes the amount of changes we make. There also is a break
#in the histograms at this value.
####----#####
data4c_corrected$days=ceiling(data4c_corrected$distance2/30)
####----#####

##Relink back to original database (which include
#the records after Sept. 2002)
newdata4=data4[substr(data4$FISHED,1,7)>"2002-09",]
```

```
newdata4$com_port=newdata4$n_com_ports=newdata4$n_areas=newdata4$distance=newdata4
$distance2=newdata4$days=NA
newdata4$trip2=newdata4$trip
####----#####
data5=rbind(data4c_corrected,newdata4)
####----#####

##-##
#-Multiday trips: only first-last reports
#Remove trips from fishers who only ever reported on the first or last day of the month
#Use only for records prior to September 2002
monthday=table(substr(data5[substr(data5$FISHED,1,7)<="2002-09",]$FISHED,6,10))
lastindex=c(31,59,91,121,152,182,213,244,274,305,335,366) #end of each month
monthday[lastindex] #check - exclude Feb. 29
firstindex=c(1,32,61,92,122,153,183,214,245,275,306,336) #first of each month
monthday[firstindex] #check
choices=c(names(monthday[lastindex]),names(monthday[firstindex])) #first and last dates of the
month

#Explore the extent to which a fisher only reports on the first or last day of the month
past_data5=unique(data5[substr(data5$FISHED,1,7)<="2002-
09",c("trip2","LICENSE","FISHED","deep7_lbs","FYEAR")]) #trips in past
past_data5$firstlast=0 #indicator if trips was first/last day of month
past_data5[which(substr(past_data5$FISHED,6,10)%in%choices),"firstlast"]=1
tab=table(past_data5$LICENSE,past_data5$FYEAR,past_data5$firstlast) #table of trips by
fishing year and license that occurred and did not occur on the first and last days of the month
per_tab=round(tab[,,2]/(tab[,,1]+tab[,,2]),2) #percent of trips by fisher-year that occurred on the
first or last day of the month

#Pull out the 646 that were 100%; only occurred on the first or last day of the month 100%
col=floor(which(per_tab==1)/nrow(per_tab)) #column occurring
row=which(per_tab==1)-nrow(per_tab)*col #row occurring
all_firstlast=data.frame("LICENSE"=NA,"FYEAR"=NA,"percent"=NA)
for(i in 1:length(which(per_tab==1))){
  all_firstlast[i,"LICENSE"]=as.numeric(dimnames(per_tab)[[1]][row[i]]) #add license # (row
names)
  all_firstlast[i,"FYEAR"]=as.numeric(dimnames(per_tab)[[2]][col[i]+1]) #add fyear (col name).
Add one to col due to flooring
  all_firstlast[i,"percent"]=per_tab[row[i],col[i]+1] #add value
}
#merge to obtain trip numbers of license-fished combinations
trips_firstlast=merge(all_firstlast,past_data5,by=c("FYEAR","LICENSE")) #trips on the first
and last days of the month from CML-fished combinations that did not record another trip
#merge with original data to obtain all records from license-fished combinations
records_firstlast=merge(data5[substr(data5$FISHED,1,7)<="2002-
09",],trips_firstlast,by=c("trip2","LICENSE","FISHED","deep7_lbs","FYEAR"))
```

```
####----#####
#exclude records from the first or last day of the month in a year
data6=data5[!data5$trip2%in%trips_firstlast$trip2,]
dim(data6)
####----#####

# --STEP03: select data representative of CPUE trends---------------------

##-##
#-Representative data: remove records from fishers that never caught weight of Deep7
####----#####
data6$fisher=gsub("[.]","",tolower(paste(data6$FNAME,data6$LNAME))) #remove punctuation
and set all names to lowercase
####----#####
checkFISH=table(data6$fisher,data6$deep7_lbs==0)
per_checkFISH=checkFISH[,1]/rowSums(checkFISH) #percent of recors with deep7
checkFISH_dataframe=data.frame("fisher"=attr(checkFISH,"dimnames")[[1]],"records"=as.num
eric(rowSums(checkFISH)),"perc_deep7"=as.numeric(per_checkFISH))
neverDeep7=checkFISH_dataframe[checkFISH_dataframe$perc_deep7==0,]
dim(neverDeep7) #1482 people never caught a Deep7
sum(neverDeep7$records) #have 7851 records
####----#####
data7=data6[!data6$fisher%in%neverDeep7$fisher,]
####----#####

##-##
#-Remove fishing independent survey trips
#Read in database of fisher, day, areas from the FI surveys
FI=read.csv("D:\\FilePath\\FI to HDAR grid_priortoFY16.csv",header=T)
FI$Date=as.Date(FI$Date,format="%m/%d/%Y")
FI$last=tolower(FI$last)
#Find FI survey trips in filtered database
FI$combo=paste(FI$Date,FI$HDAR.Grid,FI$last) #do this to combine all unique conditions> If
you cond1 & cond2.... then dont get unique conditions
a=data7[paste(data7$FISHED,data7$AREA,tolower(data7$LNAME))%in%FI$combo,]
table(a$fisher) #insure that use of last name was appropriate and only got Captains from survey
(it was)
#Check to ensure that a trip doesn't cover multiple areas (it doesn't)
dim(a[!duplicated(a[,c("FISHED","AREA","fisher")]),])
length(unique(a$trip2))
####----#####
data8=data7[!data7$trip2%in%unique(a$trip2),] #97 records, 40 trips removed
####----#####

# --STEP04: Factors for CPUE standardization---------------------

##-##
```

```
#-Add a measure of experience to each trip; cumulative number of trips
red=data8[,c("trip2","fisher","FISHED")]
ord_uniq=unique(red)[order(unique(red)$fisher,unique(red)$FISHED),] #unique fisher-trip combinations ordered by fisher
library(plyr)
num_exp=count(ord_uniq,vars="fisher") #number of unique fisher-trip combos by fisher
ord_uniq$cum_exp=unlist(lapply(num_exp$freq,FUN=seq)) #add experience value; from 1 to number of unique fisher-trip combinations per fisher
red2=join(red,ord_uniq,by=c("trip2","FISHED","fisher")) #use join so that the same order is maintained with data6
####----#####
data8$cum_exp=red2$cum_exp
####----#####

##-##
#Determine amount of Uku caught in each trip and add to data set
ukulbs=aggregate(data8[data8$SPECIES==20,]$lbs,by=list("trip2"=data8[data8$SPECIES==20,]$trip2),sum)
#merge records since many trips dont catch deep7
####----#####
data9=merge(data8,ukulbs,by="trip2",all=TRUE)
data9[is.na(data9$x),"x"]=0
dimnames(data9)[[2]][86]="uku_lbs"
####----#####

##-##
#Add quarter to the data set
#Base it on FYEAR quarter
months=as.numeric(substr(data9$FISHED,6,7))
season=months
season[months%in%c(1:3)]=3
season[months%in%c(4:6)]=4
season[months%in%c(7:9)]=1
season[months%in%c(10:12)]=2
####----#####
data9$qtr=season
####----#####

##-##
#Add wind speed and wind direction information to appropriate records.
#Wind is on a daily basis over 0.25 degree grids starting 1987-07-09
#Use getWind function which is in getWind.R
#-#
#Read in locations of management grids
library(foreign)
area_loc=read.dbf("D:\\FilePath\\Fishchart2007110807_gridpts.dbf",as.is=T)
area_locMHI=area_loc[area_loc$AREA_ID%in%valid,] #valid areas (from filtering above)
```

```
#write.dbf(area_locMHI,"D:\\FilePath\\MHI_Fishchart2007110807_gridpts.dbf")
#There are duplicate area values in Jerry and Steve's dbf file for
#area_id's 307,313,323,401,406,and 16123. Since data for only 16123
#subarea C exists in the data set, can remove all other subareas for
#16123. For all others remove the records with small length components
#which are not actually midpoitns of grids.
area_locMHI2=area_locMHI[-c(232,233,243:248,
                        282,290,292,136,140,267),]

#Assign wind information to all records within a trip
#Accounts for trips with more than one area.
#Not all locations will have viable wind data
#This takes 20 hours! <<<<<<<<<<<<<<<<<<<<<<<< NOTE <<<<<<<<<<<<<<<<<<
source("D:\\FilePath\\getWind.R")
temp=data9[data9$FISHED>="1987-07-09",]
temp[,c("speed","xdir","ydir")]=NA
u_trips2=unique(temp$trip2)
for(i in 1:length(u_trips2)){
  temp_data=temp[temp$trip2%in%u_trips2[i],]
  wind=getWind(temp_data,area_locMHI2)
  temp[temp$trip2%in%u_trips2[i],c("speed","xdir","ydir")]=

wind[match(temp[temp$trip2%in%u_trips2[i],c("AREA")],wind$AREA),c("speed","xdir","ydir
")]
  print(paste(i,"out of",length(u_trips2)))
}
#Relink with old (prior to wind data) records and form updated data set
data9_old=data9[data9$FISHED<"1987-07-09",]
data9_old[,c("speed","xdir","ydir")]=NA
####----#####
data10=rbind(data9_old,temp)
####----#####

# --OUTPUT: finalized record data set for CPUE standardization---------------------
#write.csv(data10,"D:\\FilePath\\Finalized_recordCPUE_dataset.csv",row.names=F)
```