Catch Estimation Algorithm for the Walleye Pollock *Theragra chalcogramma* Fishery and Comparison to Similar National Marine Fisheries Service Databases

Brian Battaile, Terrance J. Quinn II, David Ackley, and Galen Tromble

Reprinted from the Alaska Fishery Research Bulletin
Vol. 11 No. 1, Summer 2005
The Alaska Fisheries Research Bulletin can be found on the World Wide Web at URL:
http://www.adfg.state.ak.us/pubs/afrb/afrbhome.php

Alaska Fishery Research Bulletin 11(1):1–14. 2005.

Catch Estimation Algorithm for the Walleye Pollock *Theragra chalcogramma* Fishery and Comparison to Similar National Marine Fisheries Service Databases

Brian Battaile, Terrance J. Quinn II, David Ackley, and Galen Tromble

Abstract: As fisheries management entertains more complex objectives to ensure sustainable fisheries and ecosystems, reexamination of all aspects of data collection, data analysis, and management actions is needed. In particular, focus on fine spatial and temporal scales is becoming more common. A new spatially and temporally explicit database was constructed with this focus for the total walleye pollock (*Theragra chalcogramma*) catch in the waters off Alaska. Three sources provide information about pollock catches: the National Marine Fisheries Service observer program data, weekly processor reports to the National Marine Fisheries Service, and Alaska Department of Fish and Game fish tickets. The observer program database contains exact locations by longitude and latitude and dates. Fish tickets and weekly processor reports are much coarser in time (by cruise and week, respectively) and space (by Alaska Department of Fish and Game and federal reporting areas, respectively). Hence, obtaining spatiotemporal data at the finest scale requires maximum use of observer data. However, a significant portion of pollock catch is unobserved, so that it was necessary to combine the three data sources to provide a full accounting of catch. Comparisons were made to two National Marine Fisheries Service algorithms, the Catch By Vessel and Blend, presently used for fisheries management and analysis purposes. Estimated total catch was similar among the three systems, but the new database makes best use of the observer data and consequently is preferred for addressing fine-scale questions about pollock management.

INTRODUCTION

For fisheries management to be successful, removals from the affected fish populations must be accurately measured. Underestimation of removals from harvesting often leads to overestimation of stock abundance and underestimation of the effects of harvesting in assessment models (National Research Council [NRC] 1998). These removals include retained harvest (landings), discards, and incidental mortality. Depending on the fishery, catch information may come from at-sea observers, dockside observers, dockside landing reports (tickets, slips), and/or catcher/processor reports (NRC 2000). In many fisheries, the collection of information can be viewed as a census: records are taken from all harvesters. All the same, it is rare for all types of information to be collected from all vessels (e.g., discards, specific location of catches), so estimation of some harvest-related parameters is necessary. Finally, many fisheries have a variety of sources of information that need to be blended together. How to perform this amalgamation of information from different sources has historically received little attention in the primary fisheries literature.

There are many reasons why database work has not found its way into the primary fisheries literature. The construction of a new database and the decisions about what to include are unique to the purpose of each database with respect to the raw input material and available resources. There are some obvious general guidelines to follow (Gayanilo et al. 1997), such as accounting of all catch and avoiding redundancy; however, generic recommendations governing all situations are likely impossible. The construction of databases can even enter into the realm of professional database management instead of fisheries research. While all research and stock assessment begins with database construction, this task is often seen as a mundane and necessary evil on the road to more interesting publishable research. Literature on the subject tends to be in the form of government publications describing specific databases that may have a limited scope, such as a format for a trawl survey database (e.g. Hunter and Tremblay 1992). Other literature involves symposia or

Authors: Brain Battaile and Terrance J. Quinn II are with the Juneau Center, School of Fisheries and Ocean Sciences, University of Alaska Fairbanks, 11120 Glacier Highway, Juneau AK 99801. Email: ftbcb@uaf.edu. David Ackley and Galen Tromble are with the National Marine Fisheries Service, Alaska Region, P.O. Box 21668, Juneau, AK 99802.
Sponsorship: This publication is the result of research sponsored by Alaska Sea Grant with funds from the National Oceanic and Atmospheric Administration Office of Sea Grant, Department of Commerce, under grant no. NA 86RG0050 (project no. R/101-01) and from the University of Alaska with funds appropriated by the state.

committee publications involved in assessing a specific region's or government entity's policy on data collection and management (e.g. Sulit and Inuoe 1994) in which specific recommendations can be made for specific problems.

The collection of fisheries related information for management purposes requires significant amounts of time, effort, and money and is often the most expensive component of fisheries management. As one example, the National Marine Fisheries Service (NMFS) North Pacific Groundfish Observer Program is an extensive effort to collect data from the commercial catch for fisheries management by the North Pacific Fishery Management Council (NPFMC) (Marine Resource Assessment Group [MRAG] Americas, Inc. 2000). The need for high quality fisheries data by North American fisheries management officials is underscored by symposia and governmental reports on fisheries sampling methodology (NRC 1998, 2000; Doubleday and Rivard 1983). While basic catch quantities are the prime statistic of such data collection efforts, use of ancillary data (e.g., spatial and temporal) that are collected can be effectively utilized in increasingly complex fishery analyses. Because government agencies are constantly battling with budgetary issues, it is important to make efficient use of such data.

Little attention is given to the process between raw data collection and end user analysis. We will attempt to bridge that gap. Our objective is to compile a database accounting for all of the walleye pollock *Theragra chalcogramma* catches in the Bering Sea/Aleutian Islands and Western Gulf of Alaska (GOA) at the finest possible resolution of time, space, and catch weight. With such a database, it will be possible to explore fine-scale effects of human activities and other factors on the walleye pollock population.

We describe in detail the algorithm used to compile this database and compare the results (total estimated catch) to those of two current NMFS databases: the Blend system and the catch by vessel (CBV) system. These databases are used by NMFS for many of their management plans, stock assessments, and allocations. All three databases use the same raw data sources, and have the same basic objective of estimating total walleye pollock catch, but each has specific objectives, described below, that result in different treatment of the data sources.

MATERIALS AND METHODS

Data sources

Three sources of data were used in database construction: the NMFS observer program data, Alaska Department of Fish and Game (ADF&G) fish tickets, and weekly production/processor reports (WPR) that processing vessels provide to NMFS.

The observer program began in 1973 to observe foreign groundfish vessels operating in U.S. waters. The Magnuson Fisheries Management and Conservation Act of 1976 simultaneously created the 200 nautical mile Exclusive Economic Zone (EEZ), began the Americanization of the fishery, and established the North Pacific Fishery Management Council. Americanization was nearly complete by the late 1980s but there was no observer coverage of the domestic fleet. Consequently, the Alaska Sea Grant College Program, NMFS, and NPFMC implemented the domestic observer program, starting in 1990, to gather data to manage the wholly domestic groundfish fisheries off Alaska.

The observer program currently deploys observers based on vessel length and type of fishery operations. The following regulations apply to catcher and catcher/processor (C/P) vessels using trawl gear: vessels 125 ft and larger in overall length are required to carry an observer 100% of the time; vessels that are between 60 and 125 ft in overall length are required to carry an observer for 30% of their fishing days in each calendar quarter in which they fish for more than three days. Catcher vessels deliver to either shoreside processors (land-based plants or stationary floating processors operating in state waters) or floating processors operating in the EEZ [C/Ps or motherships (vessels that operate solely as processors) in the offshore fleet]; motherships are 100% covered by the observer program in the offshore fleet. Catcher vessels that deliver only unsorted catch from the trawl codends to processor vessels are not required to carry observers, because the hauls will be observed onboard the processor. Vessels under 60 ft do not have to carry an observer but they account for only a small percentage of the walleye pollock catch.

The haul weight recorded by observers comes from a weighing scale when the whole haul can be weighed. When direct weights are not obtained, a volumetric estimate is made, in which the volume of the catch is determined, such as the codend or bin volumes, and is then multiplied by the catch density from a sample of the catch or one prescribed by the target fishery (Alaska Fisheries Science Center [AFSC] 1999). Since 1999, most of the walleye pollock catch has been directly weighed on flow scales. Observers on C/Ps and motherships observe every haul. On smaller boats observers are unable to directly estimate every haul, in which case, estimates of total haul weight by skippers are recorded. An algorithm is then employed by NMFS, which takes that vessel's nearest observed haul in space and time and applies the species composition to the unobserved haul.

The ADF&G fish tickets are collected by the state of Alaska and are required for any groundfish landed in state waters or delivered to plants and processing vessels operating in state waters. Required information includes the date fishing began, landing date, total cruise catch weights, fishing area, and vessel information. The WPR data is collected from all processors of groundfish, independent of federal or state jurisdiction, and includes information about area fished and final product weights totaled for the week, among other data.

Each data source records fishing location in a different way. The WPRs have the coarsest scale in using the federal reporting areas (Figure 1) covering large tracts of ocean. The fish tickets use ADF&G reporting areas, which are generally 30×34.5 nmi blocks (in the eastern Bering Sea), but are subdivided near shorelines to demarcate the 3 mile state waters line and local features. The observer program records haul retrieval and deployment location by latitude and longitude coordinates to the nearest minute. The coarser reporting areas can generally be found from information on the finer scales. Hence, with ADF&G reporting areas, one can find the federal reporting area and with the latitude/longitude coordinates, one can find the ADF&G reporting area.

The NMFS Blend system

The Blend system (Figure 2) is an algorithm used by NMFS to obtain estimates of total catch from shoreside and offshore WPRs, and observer reports. For the shoreside component, WPRs are the best source of total landed catch; however, when available, observer reports are the best source for total catch including discards. Discards for unobserved vessels are estimated by multiplying known retained catch by estimates of discard rates from observed boats. Discard rates are determined by the ratio of the weight of the discarded species and the total retained groundfish weight classified by factors that define a fishery. Retained catch and estimated discards are combined for a total shoreside sector catch. For the offshore component, discards are accounted for by observers and estimated by the industry, so the Blend algorithm simply chooses between the WPR and the observer records for the corresponding

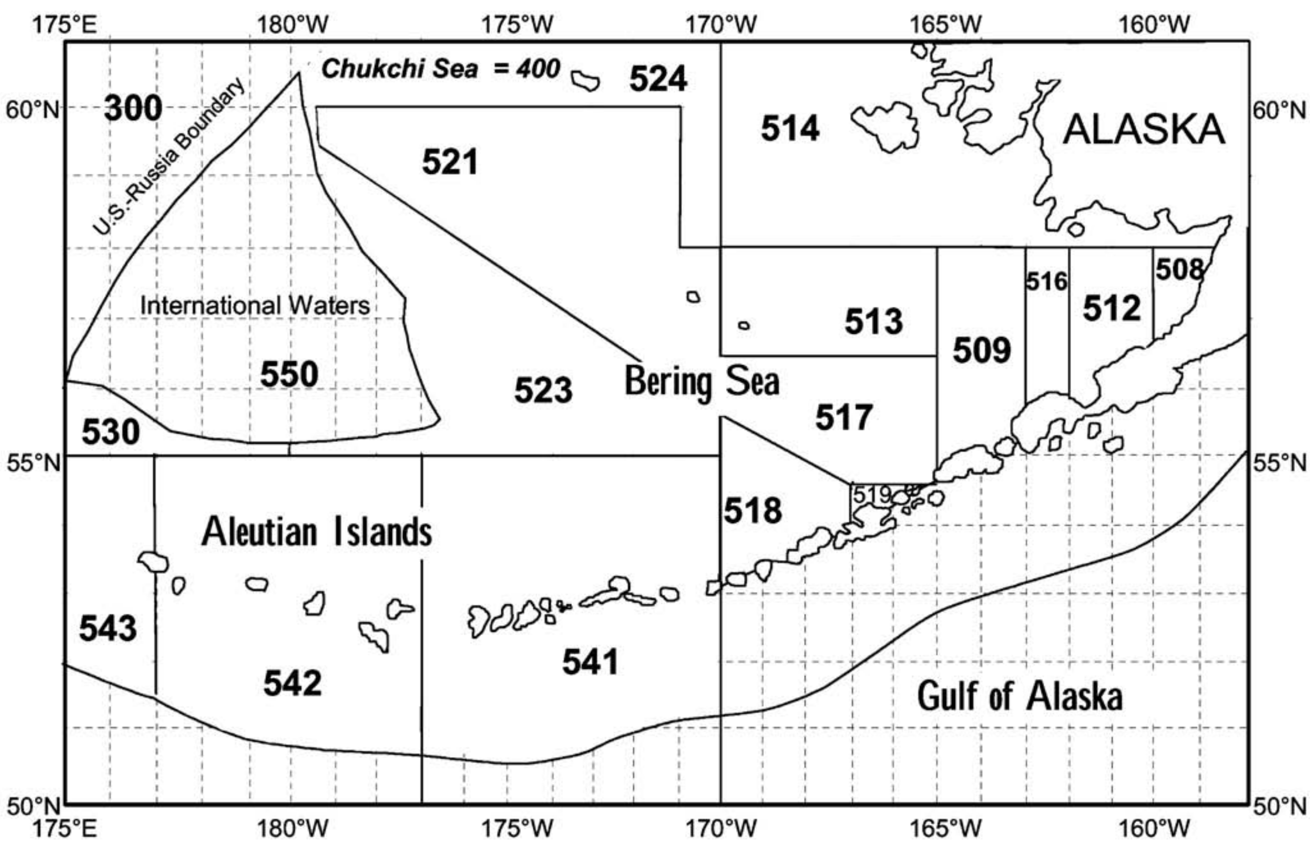

Figure 1. Map of the Bering Sea/Aleutian Islands and Gulf of Alaska detailing the federal reporting areas (NOAA 2002).

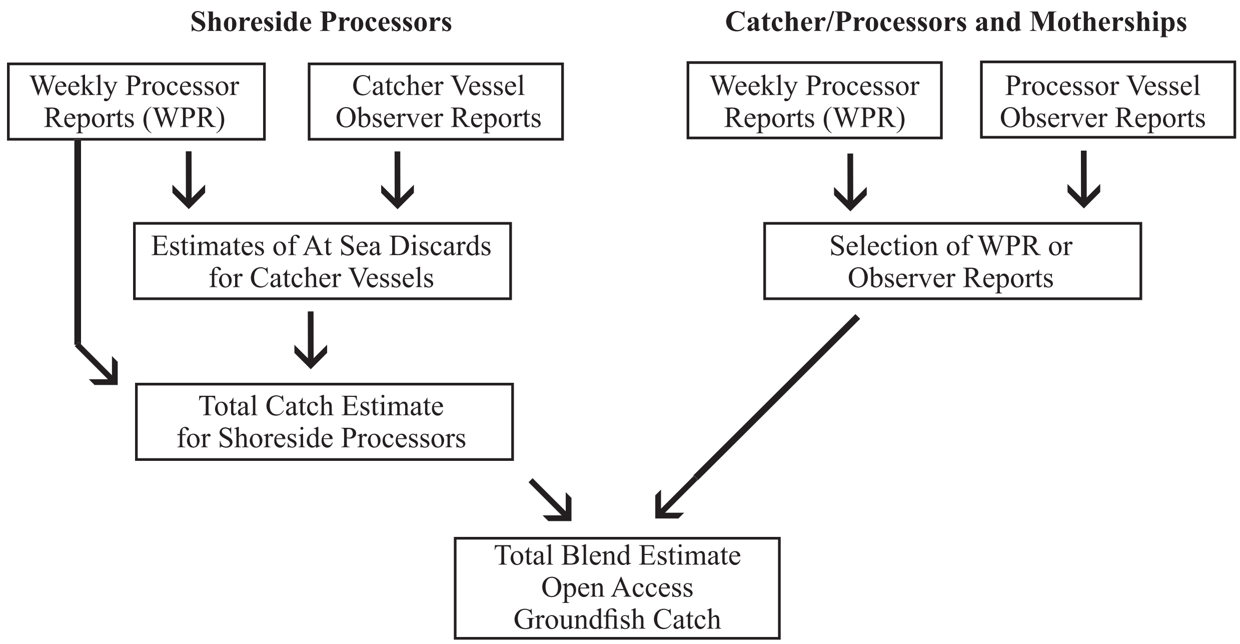

Figure 2. Flow diagram of the Blend system used by NMFS to monitor walleye pollock catch.

week to account for the total catch. The WPR record is selected in favor of observer data when: 1) the total catch numbers from WPRs and observer data are within 5% of each other, 2) the WPR is more than 30% greater in total walleye pollock than the observer total when walleye pollock is targeted, and 3) the WPR is more than 20% greater for all other groundfish species. Otherwise, the observer record is used. Rules 2 and 3, which use the WPR data, are applied when it is thought that the observer data is grossly inaccurate.

Catch By Vessel Database

With respect to walleye pollock, the Catch By Vessel (CBV) database was created to determine vessel eligibility and annual catch allocation requirements, based on historical individual vessel catch, as set forth by the American Fisheries Act passed into legislation in 1999. The two primary objectives of this database are to identify the harvesting vessel and to identify the ADF&G reporting area of catch. The ADF&G fish tickets, WPR reports, and observer data are combined to create a comprehensive database while minimizing any overlapping information. The CBV database uses a list of processors consisting of motherships and C/Ps operating in the EEZ for which fish tickets are not required. Landings delivered to these vessels (with 100% observer coverage) are represented by observer data. The ADF&G fish ticket database is the source for all other landings to processors not on this list. Fish ticket discard information is not included. The WPR data are selected for deliveries to all C/Ps on the list less than 125 ft in length.

Spatial and Temporal Scale Database

The primary data source for the spatial and temporal scale (STS) database is haul information from the NMFS observer program. For those boats or trips not covered by the observer program, ADF&G fish tickets from shoreside processors (for unobserved vessels delivering to shoreside processors) and WPRs from at sea processors fill in remaining data gaps. The observer database is our preferred source because it records catch weight by haul, haul time to the nearest minute from the ship's log book, and location to the nearest minute in latitude and longitude. The ADF&G fish tickets record landings (generally weighed on scales), fishing location by ADF&G reporting area, the date fishing began, and occasionally estimates of discards, but does not record a more specific measure of effort such as haul time or number of hauls. The WPRs also lack effort data, report location using the coarse federal reporting areas only, and estimate weekly catch from finished product, making it the least desirable data source. Hence, any duplicate data from the three sources is always represented by the observer program in the STS.

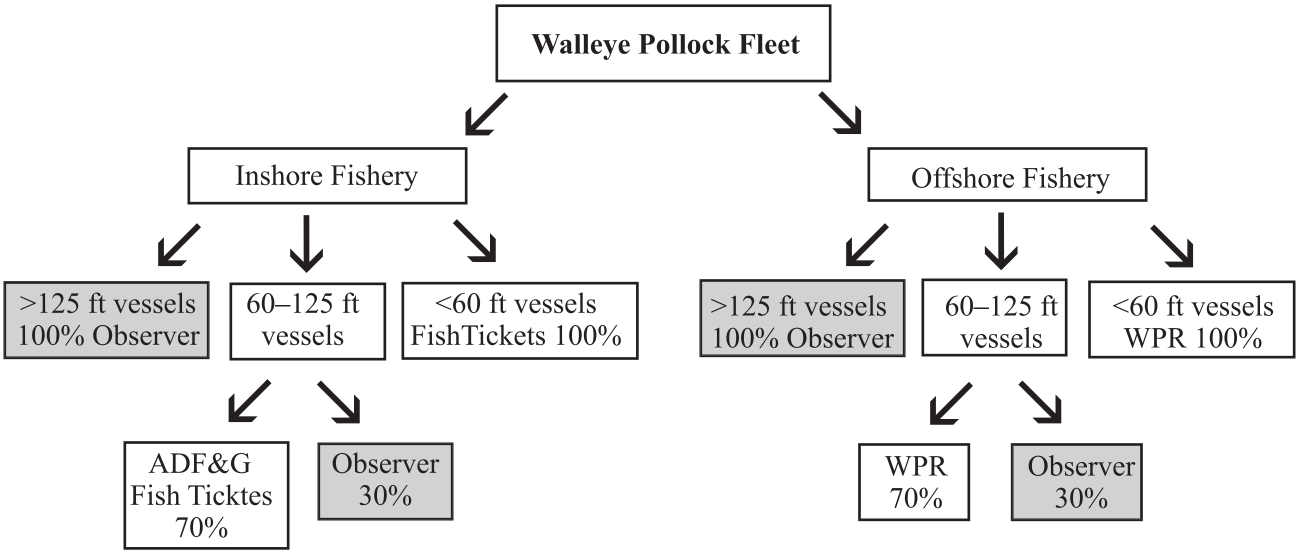

Figure 3. Flow diagram for data sources used in the catch estimation algorithm: data sources include observer coverage (Observer), Alaska Department of Fish and Game (ADF&G) fish tickets and weekly processor reports (WPR). Approximate percentages of fishing trips covered by each data source in each sector (Inshore, Offshore) are given.

Data algorithm for the STS

Figure 3 is a flow chart of expected percentages of data from the three source databases contributing to the STS. Theoretically, catches from 70% of the fishing days of 60–125 ft vessels in the inshore sector are recorded only with ADF&G fish tickets. The 30% (approximately) of fish tickets also covered by the observer program were identified by vessel ID number and overlapping dates, and matched as follows. All observed hauls should have a corresponding fish ticket. Fish tickets record the fishing start date and the landing date while observer records show the date a haul takes place. Hauls for a specific vessel, recorded by the observer program, falling within or on fish ticket dates for the same vessel were removed by ordering the two sets of records by date on a computer spreadsheet. It was assumed that observers were present for a complete trip. Like the CBV, fish ticket discard information was not included.

Weekly processor reports for the offshore sector report a total catch estimate using an algorithm (performed by NMFS) to extrapolate weights of processed product to unprocessed catch weights. The WPR data were used for the offshore sector when there were gaps in the observer record. Data gaps could occur for unobserved 60–125 ft C/Ps (70% of the time) and in rare instances, for observer records determined by NMFS to be flawed.

The process of using WPR data was as follows. First, if all days within the WPR record were observed, or if observed catch exceeded WPR catch, then the WPR data were eliminated. Second, if observer coverage was incomplete for a week and when there was a positive difference between the haul estimates (WPR Observer), the difference was included in the STS. If WPRs indicate that fishing occurred in more than one federal area for the week, then the catch was partitioned over those areas. Finally, for weeks with no observer coverage, WPR haul weights were simply recorded for the STS.

Comparison between databases

Very little by means of formal statistical testing is applicable here as we are making a census of the catch and defining a selection process. We do, however, compare totals between the STS, CBV and Blend database from the observer, fish tickets and WPR sources. We further stratify the totals by ADF&G and federal reporting areas, vessel sizes, and processor categories to examine characteristics of the pollock fishery and how the information in the STS, CBV and Blend databases reflect the fishery.

RESULTS

Tables 1, 2, and 3 show the relative contributions of the three raw data sources, and the final total weights for the years 1995–1999 for the Blend, CBV and STS databases, respectively. The Blend system relies on

Table 1. Walleye pollock catch in weight (metric tons) by database source for the Bering Sea (BS), Aleutian Islands (AI) and the Gulf of Alaska (GOA) fisheries using the Blend system.

Source	1995	1996	1997	1998	1999
BS					
Observer records	785,275	732,745	648,864	515,454	549,014
WPR records	424,892	384,295	456,971	566,174	429,738
Estimated discards	15,076	13,803	18,758	2,247	10,444
BS Total	1,225,243	1,130,842	1,124,592	1,083,875	989,196
AI					
Observer records	71,082	75,598	12,948	18,290	1,087
WPR records	33,088	15,698	12,976	22,051	48
Estimated discards	90	201	16	1	385
AI Total	104,260	91,497	25,940	40,342	1,520
GOA					
Observer records	2,056	2,241	1,613	189	374
WPR records	64,857	47,048	85,211	123,965	93,438
Estimated discards	5,705	1,974	3,260	1,306	1,825
GOA Total	72,618	51,263	90,085	125,460	95,637
BSAI Total	1,329,503	1,222,339	1,150,532	1,124,217	990,717
BSAI-GOA Total	1,402,122	1,273,602	1,240,617	1,249,677	1,086,354

Table 2. Walleye pollock catch in weight (metric tons) by database source for the Bering Sea (BS), Aleutian Islands (AI), and Gulf of Alaska (GOA) fisheries using the catch by vessel (CBV) system. The total BSAI weight from the CBV system compared to the Blend system is shown as % BSAI of Blend.

Source	1995	1996	1997	1998	1999
BS					
Fish Tickets	410,412	389,228	352,756	364,900	433,545
Observer records	839,315	774,921	739,223	723,284	543,763
WPR records	34	203	69	1,022	1,237
BS Total	1,249,761	1,164,352	1,092,048	1,089,206	978,544
AI					
Fish Tickets	17,367	11,043	8,257	7,616	1
Observer records	45,518	18,151	18,742	15,937	737
WPR records	0	0	0	3	0
AI Total	62,885	29,193	27,000	23,555	738
GOA					
Fish Tickets	66,637	49,244	87,102	131,941	94,957
Observer records	1,618	2,322	645	269	405
WPR records	0	45		43	79
GOA Total	68,256	51,611	87,747	132,253	95,442
BSAI Total	1,312,646	1,193,545	1,119,048	1,112,761	979,282
BSAI-GOA Total	1,380,902	1,245,156	1,206,795	1,245,014	1,074,724
% BSAI of Blend	98.88%	97.76%	97.45%	98.82%	98.92%

observer records for approximately 50–70% of the BSAI data over these five years, with WPRs supplying the remainder (Table 1). In contrast, the WPRs are the dominant data source in the GOA (approximately 90%), because the smaller vessels of the GOA have less observer coverage.

The CBV system relies on observer data primarily with the remainder supplied from ADF&G fish tickets (Table 2). Clearly, WPRs contribute very little to the CBV. As in the Blend system, some differences occur between regions, with the GOA region heavily dependent on fish ticket data for the CBV.

The STS system relies on observer data for the majority of its source data in the BSAI (approximately 90%) and an insignificant amount on the WPR data (Table 3). The percentages of observer coverage for the STS are similar to that expected from Figure 3 for the medium-sized vessels of the inshore sector (Table 4) with ~70–80% from fish tickets and ~20–30% from observer records. As in the CBV, fish tickets supply the majority of the data for the STS in the GOA.

The total walleye pollock catch weight for each year from the STS database is consistently larger than that for the CBV by ~2.50% for 1995–1998 and

Table 3. Walleye pollock catch in weight (metric tons) by database source for the Bering Sea (BS), Aleutian Islands (AI), and Gulf of Alaska (GOA) fisheries using the spatial and temporal scale (STS) system. The total weight from the STS system compared to the catch by vessel (CBV) system is shown as % BSAI of CBV; the total weight from the STS system compared to the Blend system is shown as % BSAI of Blend.

Source	1995	1996	1997	1998	1999
BS					
Fish Tickets	123,634	130,050	109,550	106,785	140,435
Observer records	1,157,131	1,061,254	1,005,102	1,012,020	888,747
WPR records	1,664	2,217	4,205	501	720
BS Total	1,282,430	1,193,521	1,118,857	1,119,306	1,029,902
AI					
Fish Tickets	2,164	949	882	490	
Observer records	63,386	27,643	26,524	22,226	739
WPR records	2	0	0	0	0
AI Total	65,552	28,592	27,406	22,716	739
GOA					
Fish Tickets	37,343	33,686	58,623	85,661	67,849
Observer records	37,352	22,286	35,841	48,267	32,267
WPR records	792	946	513	66	97
GOA Total	75,486	56,918	94,976	133,994	100,213
BSAI	1,347,982	1,222,113	1,146,263	1,142,022	1,030,641
BSAI-GOA	1,423,468	1,279,032	1,241,239	1,276,016	1,130,854
% BSAI of CBV	102.69%	102.39%	102.43%	102.63%	105.24%
% BSAI of Blend	101.39%	99.98%	99.63%	101.58%	104.03%

Table 4. Total weight (in metric tons) of walleye pollock by medium-sized vessels (66–125 ft) accounted for by fish tickets in the spatial and temporal scale (STS), and the total catch by medium-sized vessels. The expected 70% coverage by fish tickets in 1995–1996, and increase to 81% coverage by 1999 are shown.

STS medium vessels	1995	1996	1997	1998	1999
Fish Tickets	150,044	150,632	150,097	168,434	193,513
Total	221,186	205,628	197,343	220,343	238,149
Fish Tickets/Total	68%	73%	76%	76%	81%

by 5.25% for 1999 (Table 3). The STS total catch is within –0.50% to 1.50% of that for the Blend system for 1995–1998 and greater than that for the Blend by 4.25% in 1999.

Figure 4 graphically compares the relative contributions of catch from the source databases used by each database. The contributions for the BS and AI regions are relatively similar, but those from the GOA are quite different due to its heavier dependence on a shore-based fleet. The STS clearly contains the largest percentage of observer data, the CBV contains the largest percentage of fish ticket data, and the Blend contains the largest percentage of WPR data.

In order to examine the consistency among the databases, the proportions of catch by vessel size (small, medium, and large) for the STS, CBV, and Blend databases are graphically compared for each region and year (Figure 5). The CBV and STS databases on the middle and outer rings show nearly identical relative percentages of catch by vessel size indicating consistency through the source databases. The Blend database uses shore-based processor WPRs with no associated catcher-vessel length data; hence no comparison to the other databases can be made.

The percentages of observer data used in the CBV and STS databases change across ADF&G areas (Figure 6; the Blend database could not be included because its WPR data do not have location information at the ADF&G area level). In general, a lower percentage comes from observer records near shore and increases as the distance from the shore increases, and this trend is more pronounced in the CBV relative to the STS. Some northern areas show the CBV with a slightly greater percentage of observer data than the STS. This does not indicate that the CBV included observer data that the STS did not; instead, the STS algorithm included a small number (1–3) of fish tickets that the CBV did not. This could happen if a boat designated as "offshore" offloaded onshore. Similarly, a federal stratification shows the same general trend (Figure 7, including the Blend database) and also indicates the large differences between the percentage of observer data between the Blend and STS.

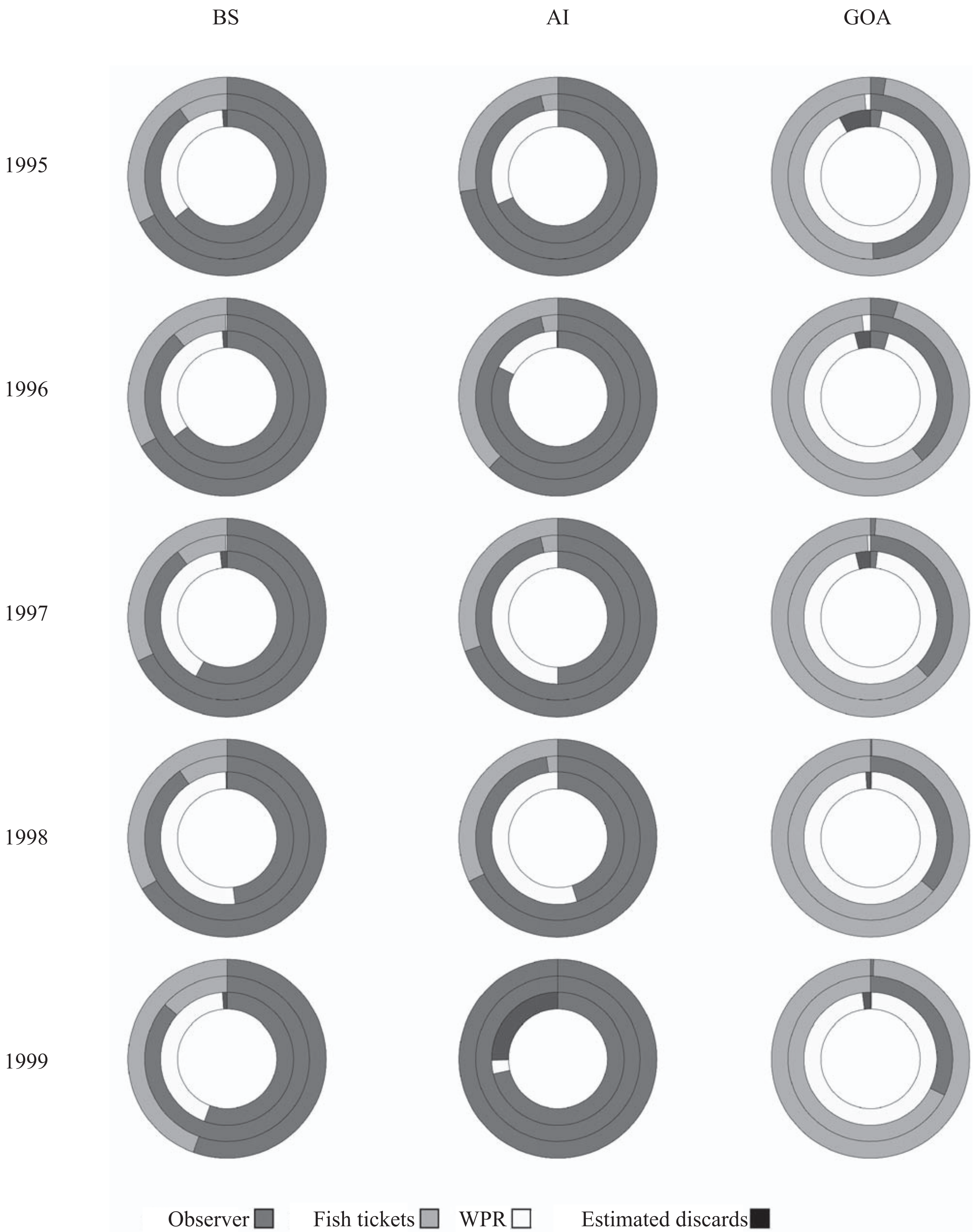

Figure 4. Estimated relative proportion of walleye pollock catch and discards by source database categories using the Blend (inner ring), spatial and temporal scale (STS, middle ring) and catch by vessel (CBV, outer ring) databases.

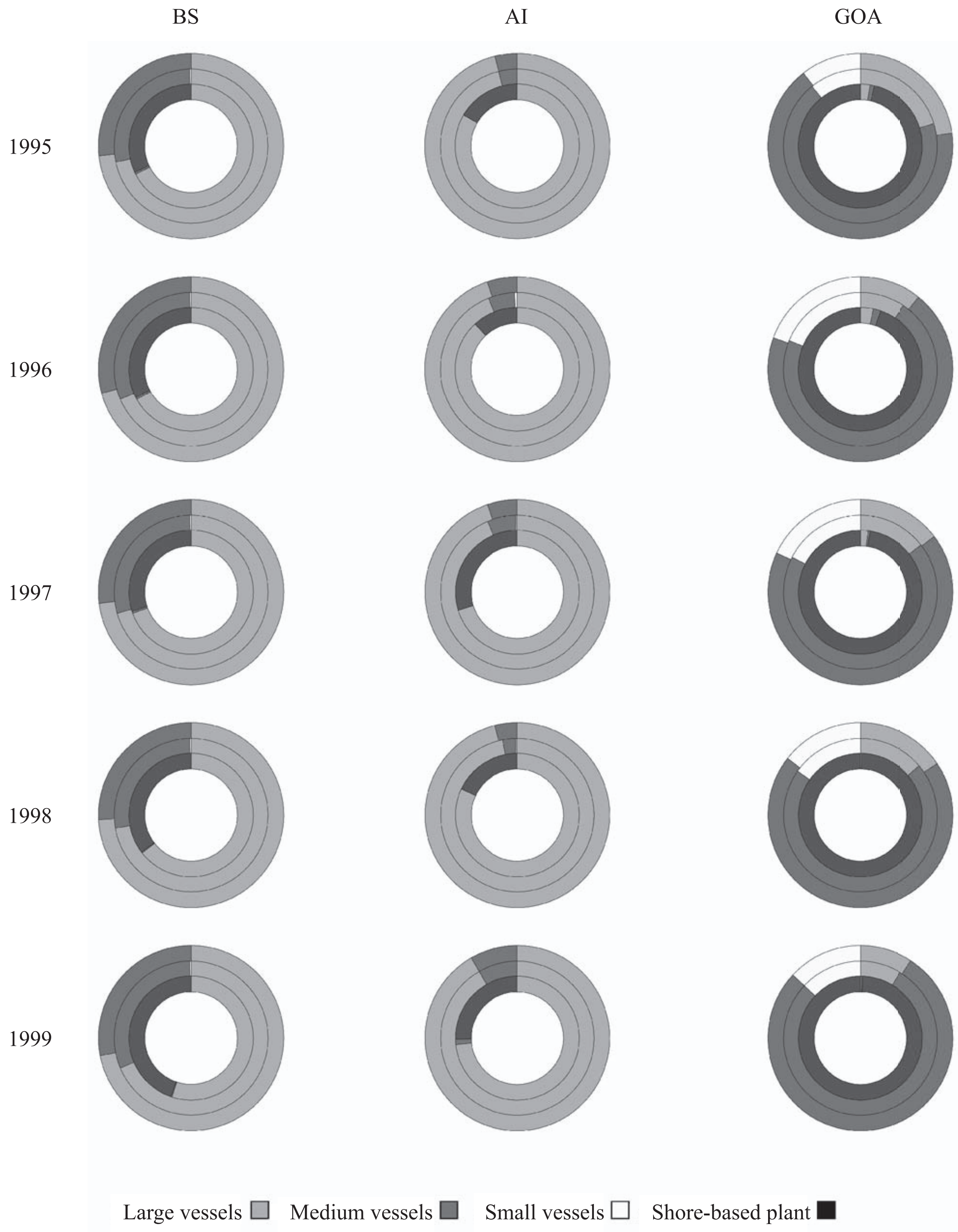

Figure 5. Estimated relative proportion of walleye pollock catch and discards by vessel size or processor categories using the Blend (inner ring), spatial and temporal scale (STS, middle ring), and catch by vessel (CBV, outer ring) databases.

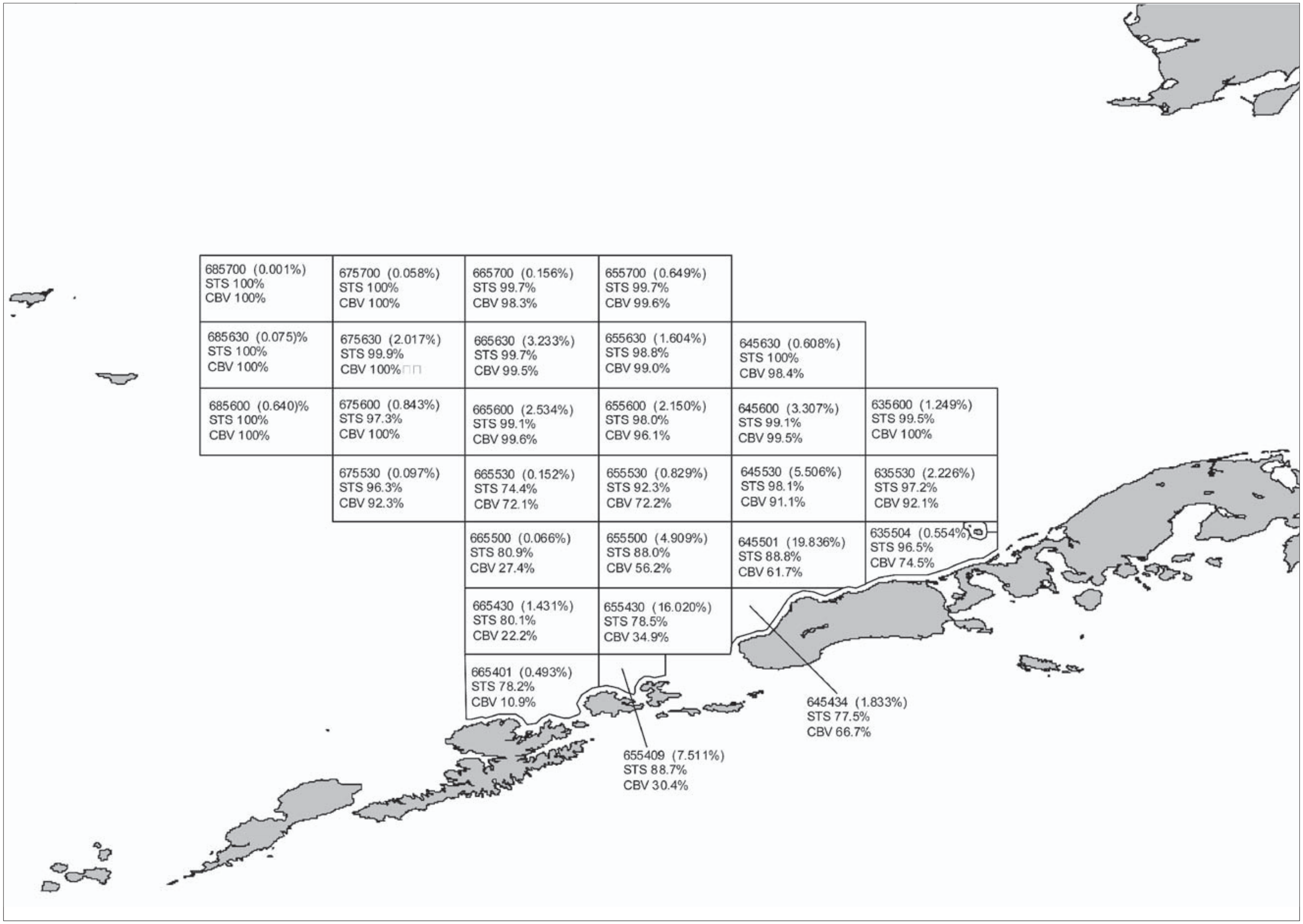

Figure 6. Relative percentages of observer program data for the catch by vessel (CBV) and spatial and temporal scale (STS) databases aggregated by Alaska Department of Fish & Game reporting areas. The 6-digit number is the ADF&G reporting area code, followed by the percentage of the total database by weight assigned to that area. The STS and CBV percentages indicate the percentage of data in that area that comes from observer program data.

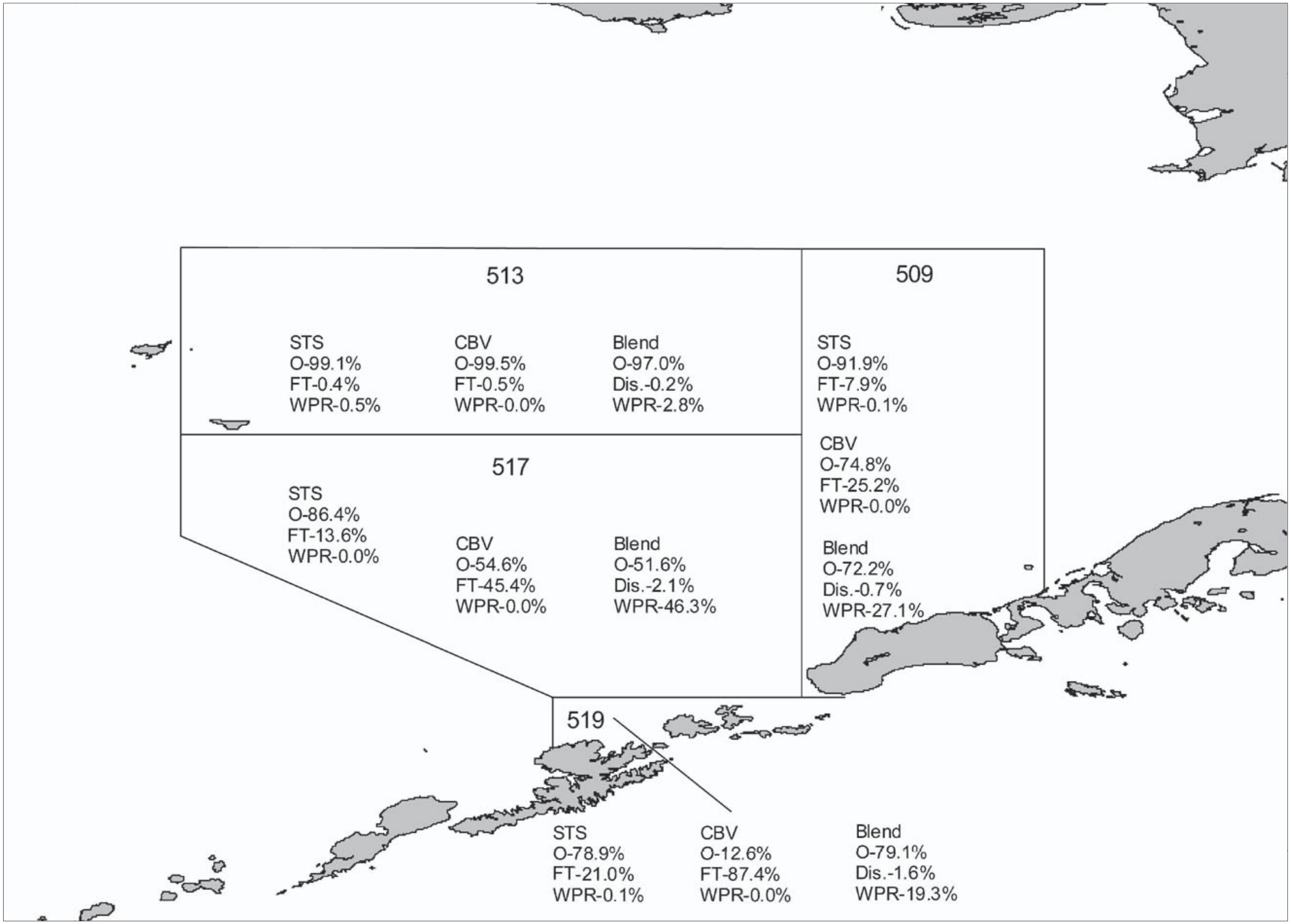

Figure 7. Relative percentages of observer program data for the catch by vessel (CBV) and spatial and temporal scale (STS) and Blend databases aggregated by Federal reporting areas. The 3-digit number is the Federal reporting area code. The percentages indicate the percentage of data in that area that comes from observer (O) program data, fish tickets (FT), estimated discards (Dis.) and weekly processor reports (WPR).

Table 5. Differences among observer data and weekly processor reports (WPRs) and fish tickets (FT) for 1995. Weights are in kilograms. Observer weights are always larger for the total comparison of each database. Complete coverage indicates WPR weeks that had observer coverage through the entire week, incomplete coverage indicates observers were present for part of the week. Inseason hauls indicate hauls made within the prescribed pollock season while off-season hauls were made outside of the pollock season.

	Fish ticket		WPR	
	Off-season Hauls	Inseason Hauls	Complete Coverage	Incomplete Coverage
Number of FT or WPR records	147	420	70	119
Total FT or WPR weight	604,000	74,749,000	1,738,000	2,055,000
Corresponding observer weight	2,684,000	77,133,000	1,766,000	2,091,000
Total difference	2,080,000	3,384,000	28,000	36,000
Percent difference of observer to FT or WPR	444%	4.4%	1.6%	1.7%
Observer haul average	18,259	143,370	31,530	21,117
Mean difference between individual observer and WPR or FT records	14,146	8,057	614	2,529
CV of mean difference	7.2%	23%	780%	96%
Ratio (observer:WPR or FT) geometric mean	27.07:1	1.05:1	1.10:1	1.33:1
CV of geometric mean	38%	38%	90%	1105%

Unexpectedly large differences exist between WPR and observer, and fish ticket and observer records. For the 1995 data, we made a comparison of matched fish tickets and observer data and matched WPR and observer data (Table 5). The fish ticket and observer matched data are split into 2 groups, those within the walleye pollock season and those outside the walleye pollock season. During the walleye pollock season, the weights from fish tickets are relatively close overall (96%) to the observer data and individually, the geometric mean of the ratio of the individual fish tickets and their matched hauls is 0.95, quite similar to the expected ratio of 1. The off-season fish tickets are much different in total (444%) and individually with a geometric mean of the ratios of 27:1. The WPR data shows little difference in total weight when compared to the observer program but differences between individual records and their matched observer records are quite variable; coefficients of variation for WPRs with complete observer coverage and those with incomplete observer coverage are 780% and 96%, respectively.

DISCUSSION

The objective of the STS was to compile a database accounting for the total walleye pollock catch while maximizing spatial and temporal detail useful for the detailed fisheries analysis. As mentioned earlier, data from the observer program accomplishes this objective better than the other 2 data sources; hence, maximizing data from the observer source was the primary objective. Nearly 90% of the documented walleye pollock catch in the STS was taken from the observer program, 10% from the ADF&G fish tickets, while the WPR contribution represented less than 1% (Table 3, Figure 4). Nearly 100% of the database includes haul location to the level of ADF&G reporting area and nearly 90% includes effort data as haul duration to the minute and a spatial resolution for fishing location to the nearest minute of longitude and latitude. The CBV is made up of much more fish ticket data than the STS, and the Blend database has considerable percentages of WPR data, particularly in 1998 (Tables 1, 2, 3 and Figure 4).

Vessel size and sector differences affect the relative percentages of observer, fish ticket and WPR data in a particular area. Very few vessels are under 125 ft in the offshore sector; hence most of these vessels were covered by the observer program. Thus, very little WPR data were incorporated in the CBV and STS databases. The GOA has relatively few larger boats, so less observer program data are utilized. The spatial differences in source data percentages between the CBV and STS shown in Figure 6 arise directly from the interaction between vessel size and observer coverage. These spatial differences show that no one data source alone can be used as a relative index of catch by area, because fish tickets cover a disproportionate number of nearshore small vessels that are not observed.

In comparing the three methods, the STS system is quite close to the Blend and the CBV in total catch. The CBV has slightly lower percentages for total catch than the STS in the BSAI (Table 3). The 1999 STS totals are greater than the CBV and Blend by a few percentage points relative to the other years. In 1999, major changes in the walleye pollock fishery occurred due to the American Fisheries Act. It is also evident that the CBV has a larger percentage from fish ticket data in 1999 than in previous years, as did the STS.

The STS is approximately 5% larger in both the GOA and BS areas. Catch in the AI is practically the same, but walleye pollock is only a bycatch component in this area in 1999. Also in 1999, the greatest percentage (81%) of catch (from fish tickets) of shore-based medium-sized vessels is about 10% larger than expected from Figure 3 due to new catch allocation from the American Fisheries Act.

The fundamental differences between the Blend, CBV, and STS stem primarily from main organizational categories under which the data are collected. The CBV is vessel oriented, hence the "catch by vessel" name. The Blend is processor oriented, hence WPR data are most prevalent. The STS is observer oriented and more similar to the CBV than to the WPR. With the CBV, vessel information was obtained such that one source of information, be it observer, fish ticket or WPR, could completely cover a category of vessels. Thus, the CBV avoided the matching of data that occurred in the STS and hence lessened the problems of date and time inconsistencies between observer, fish ticket, and WPR databases. Because the CBV and STS are most similar, comparison of the CBV database to the STS is perhaps most telling of the accuracy of the final haul weights found in the STS. In comparing the BSAI data, the greatest difference between the two databases is 5% (Table 5).

Differences between duplicate data reported in the observer data, WPR reports, and fish tickets, and the differences in how these are dealt with among the Blend, CBV and STS databases could explain the different yearly totals found among databases (Table 3). In compiling the STS, subsequent fishing trips would often have the same landing and start dates in the fish tickets, making it difficult to match observer hauls to a particular fish ticket on those dates. In addition, trip start dates have been found to be prone to error because they were recorded at the time of landing. Fish tickets and observer data covering the same hauls never report the same weights, although observer and fish ticket data for inseason large walleye pollock hauls (in the 10s to 100s of thousands of kilograms) were more similar than off-season hauls (Table 5). Fortunately, off-season catches are orders of magnitude smaller. Such problems could lead to differences between 2 databases that use different matching algorithms and different proportions of the fish ticket and observer data. The latter situation is especially problematic if either the fish tickets or observers tended to record larger amounts for the same haul. The inaccuracies of the fish ticket start and landing dates are probably the largest source of error for the STS. Fortunately, the inconsistencies between observer data and fish tickets occur in only about 10% of the fish tickets, which comprise about 10% of the STS, resulting in only 1% of the STS database with potential errors, a relatively small fraction. Hence, the errors are unlikely to cause major problems when using the STS for data analysis.

The large differences between off-season WPR and observer reports are somewhat disturbing. While off-season catch makes up a minute percentage of the total, its inaccuracies could have a significant impact on analysis investigating off-season catch. In such a case, further consideration should be given to determining which database is most accurate.

Discards from individual ADF&G fish tickets have not been included in the STS. They are recorded on the fish ticket as estimated by the fisher but not accurately measured during the fishing trip. For this reason, they have not been included here, a decision also reached in compiling the CBV dataset. One reason why the total catch from the STS is larger than that for the CBV may be because the CBV has a larger percentage of data from fish tickets and discards from these trips were excluded, whereas discards are included in the corresponding observer data. How discards are treated in the STS should be closely considered when using it for data analyses, as it may be significant. Average discard rates in the eastern Bering Sea from 1995–1997 and 1998–1999 averaged 7.3% and 2.25%, respectively (Ianelli, et al. 2003).

CONCLUSION

The relative agreement of the three databases lends credence to the legitimacy of all three. While differing in their primary applications and hence, source data ratios, all three accurately depict the exploitation of the walleye pollock fishing fleet. The STS has significantly improved spatial and temporal resolution compared to the Blend and CBV databases, while still providing consistent annual total walleye pollock catch estimates The STS also uses the largest percentages of observer data, which are collected contemporaneously with the catching process. This study has shown that increased awareness of data management can maximize the utility (scientifically and financially) of available data.

Since 2003, the region has used a catch accounting system that took advantage of reporting, especially from shoreside processors, at a more detailed level than had previously been available for the Blend algorithm. Most of the data is now received electronically and processed in a timely manner through the catch accounting system for quota monitoring on an Oracle database. Electronic landing reports supplied by processors are currently the source of catch data for the vast majority all groundfish landed at shoreside locations. These data are at the individual vessel and trip level, and equivalent to fish tickets collected by the ADF&G. A small portion of shoreside processors not currently using the electronic reporting software are also required to submit weekly processor reports and this is the primary data source from those plants.

LITERATURE CITED

AFSC (Alaska Fisheries Science Center). 1999. North Pacific groundfish observer manual. North Pacific Groundfish Observer Program. Seattle, Washington.

Department of Commerce. 1997. Fisheries of the exclusive economic zone off Alaska; Improved retention/improved utilization. Federal Register: December 3, 1997. 62(232): 63880–63891.

Doubleday, W. G., and D. Rivard, editors. 1983. Sampling commercial catches of marine fish and invertebrates. Canadian Special Publication of Fisheries and Aquatic Sciences 66.

Gayanilo, R. C., Jr., T. Stromme, and D. Pauly. 1997. Toward a generic trawl survey database management system. Pages 116–132 *in* G. Silvestre and D. Pauly, editors. Status and management of tropical coastal fisheries in Asia. ICLARM Conference Procedings 53.

Hunter, C. M. and M. J. Tremblay. 1992. A database for catch and length composition data related to the Scotia-Fundy lobster fishery. Canadian Manuscript Report of Fisheries and Aquatic Sciences 2149.

Ianelli, J. N., S. Barbeaux, G. Walters, and N. Williamson. 2003. Bering Sea-Aleutian Islands walleye pollock assessment for 2003 *in* Stock assessment and fishery evaluation report for the groundfish resources of the Bering Sea/Aleutian Islands regions. North Pacific Fishery Management Council Anchorage, AK, 1:1–126.

MRAG (Marine Resources Assessment Group) Americas, Inc. 2000. Independent review of the North Pacific Groundfish Observer Program. Tampa Bay, Florida.

NRC (National Research Council). 1998. Improving Fish Stock Assessments. National Academy Press, Washington DC.

NRC (National Research Council). 2000. Improving the collection, management, and use of marine fisheries data. National Academy Press, Washington DC.

NMFS (National Marine Fisheries Service). 2000. Endangered Species Act-section 7 consultation, biological opinion and incidental take statement. November 30, 2000. National Marine Fisheries Service, Juneau, Alaska.

NOAA (National Oceanographic and Atmospheric Administration). 2002. 2003 Record keeping and reporting (R&R)–Figures, BSAI Statistical and Reporting Areas, 2002, National Marine Fisheries Service. http://www.fakr.noaa.gov/rr/figures.htm Accessed March 20, 2003.

Seber, G. A. F. 1982. The estimation of animal abundance and related parameters, Second edition, Griffin, London.

Sulit, V. R., and K. Inoue. 1994. Status of fishery information and statistics in Asia. Proceedings of the regional workshop on fishery information and statistics in Asia. Southeast Asian Fisheries Development Center (SEAFDEC) Secretariat, Bangkok, Thailand Vol. II: Report of the Workshop.