NOAA-NMFS-NWFSC TM-31: Data Collection -- Groundfish (cont):

CHAPTER 6
PACIFIC FISHERIES INFORMATION NETWORK

William P. Daspit
Pacific States Marine Fisheries Commission, Pacific Fisheries Information Network
Building 4 - Room 2066, 7600 Sand Point Way NE, Seattle, WA 98115, USA
Paul R. Crone
National Oceanic and Atmospheric Administration, National Marine Fisheries Service
Northwest Fisheries Science Center, Fishery Analysis and Monitoring Division
Hatfield Marine Science Center, 2030 S. Marine Science Drive
Newport, OR 97365, USA
David B. Sampson
Oregon State University, Coastal Oregon Marine Experiment Station and
Department of Fisheries and Wildlife, Hatfield Marine Science Center
2030 S. Marine Science Drive, Newport, OR 97365, USA

6.1 Introduction

Since 1974, the Pacific States Marine Fisheries Commission (PSMFC) has worked actively with its member states and federal agencies to improve the quality and timeliness of data collection, processing, and analysis of fishery information and to produce data summaries required for regional conservation and management purposes. This effort was recommended initially by leaders from the albacore fishing industry, who urged management agencies to organize coastwide databases for fish landings, fishing effort, and characteristics of fishing vessels for all fisheries of the U.S. Pacific coast. These leaders recognized that simple summation of the results generated independently by each state could lead to serious misconceptions regarding the status of a fishery, because of the different sampling and analytical methods used by the individual states. In particular, highly mobile fisheries that span state boundaries would greatly benefit from a coastwide database that was accurately maintained. The landings data that existed prior to 1974 (see section 1.2) was insufficient for coastwide in-season quota monitoring.

This coastwide data coordination and consolidation effort received major impetus from enactment of the Magnuson Fisheries Conservation and Management Act of 1976, which established Regional Fishery Management Councils charged with managing fishery resources as geographical units throughout the range of the species

on the basis of the best available scientific information. It was clear that regionally comprehensive and coherent fisheries data were needed on a timely basis to provide the information required by the Regional Fishery Management Councils.

Regional fisheries data coordination requires effective cooperation and mutually supportive interactions among state fisheries agencies, which on the Pacific coast collect all commercial catch statistics from domestic fishers who land their catch at shoreside ports in the United States, and among Pacific-area National Marine Fisheries Service (NMFS) Regions and Centers, which are responsible for collection of all data for fisheries that operate in the U.S. Exclusive Economic Zone. To assure effective communication and cooperation among those state and federal entities, the Pacific area has been served since 1974 by a sequence of regional coordinating committees comprised of representatives from the participating fishery agencies.

First, there was the Albacore Coordination Committee and its Data System Task Group, which was superceded by a NMFS-sponsored committee known as the Coastwide Data Task Force. The Committee on Goals and Guidelines for Regional Fisheries Data Collection was then established and restructured in 1980 as the Pacific Coast Fisheries Data Committee, which remains the name of the regional coordinating committee today.

The Data Committee consists of 13 members appointed by the directors of the following participating agencies: Alaska Department of Fish and Game (ADFG), California Department of Fish and Game (CDFG), Idaho Department of Fish and Game (IDFG), Oregon Department of Fish and Wildlife (ODFW), Washington Department of Fish and Wildlife (WDFW), Pacific Fishery Management Council (PFMC), North Pacific Fishery Management Council (NPFMC), six Centers and Regions of the NMFS, and PSMFC. The member appointed by the Alaska Fisheries Science Center (AFSC) also represents the Northwest Fisheries Science Center (NWFSC).

The Data Committee was chartered in 1980 with four stated goals.

1. Implement and manage a Pacific Fisheries Information Network (PacFIN), which is responsible for aggregating detailed and summarized state and federal fisheries data that are used by resource managers and associated fishery-related agencies.

2. Provide data-management consultation and technical advice to the Council's Management Teams and participating agencies upon request.

3. Establish priorities and coordinate plans to improve the efficiency and timeliness of data acquisition and delivery with a minimum of unnecessary duplication.

4. Promote the development and implementation of coastwide data-collection standards to facilitate aggregation of fisheries data within the PacFIN system.

The overall PacFIN system is expansive, and many of the intricacies involved in accessing, retrieving, and interpreting the data that reside in the system are beyond the scope of this document. This chapter focuses on the procedures and components of the PacFIN system that are relevant to the groundfish fishery data collected and submitted by the three state fishery agencies of the U.S. Pacific coast: WDFW, ODFW, and CDFG. Brief descriptions of related information are also presented to complement primary areas of discussion. Information presented in this chapter summarizes an extensive description of the PacFIN system that is available through the PacFIN office in Seattle (Daspit 1996).

6.2 PacFIN System, 1981-87

6.2.1 Groundfish Data Collected by State Fishery Departments

In February 1981, the Data Committee hired the system designer/manager to design and implement the PacFIN system. Prior to this, the Data Committee had met on eight occasions over a two-year span and produced an initial requirements document that became the starting point for system development. One requirement was that the system would be operational within six months. Others were that input data would be provided to the central database on the 15th of each month, that data for the month ending 15 days earlier would be 90% complete, and that all earlier months would be more complete than the most recent month. Issues regarding the confidentiality of fishery-related data were discussed at the February 1981 meeting of the Data Committee. Discussions focused on developing protocols for data acquisition that met all legal requirements and that allowed researchers and managers to obtain needed information easily. The consensus of the Committee was to avoid the confidentiality issue by specifying a system that required only data that were aggregated to some reasonable and useful higher level. Individual fish tickets and vessel registration records were specifically ruled out of consideration as possible information to be included in the central database. The Groundfish Management Team (GMT) of the PFMC produced specifications for two initial reports that addressed primary retrieval requirements. One report presented monthly catch estimates by species and International North Pacific Fisheries Commission (INPFC) area, and another report provided monthly catch estimates by species and data source (i.e., agency providing fishery-related data), including foreign countries and joint-venture (JV) operations. A

system specification was produced in May 1981 and the initial implementation of the PacFIN system was operational in October 1981. The system was developed on a Burroughs B7800 computer that was owned, operated, and maintained by the Office of Fishery Information Systems of the Northwest and Alaska Fisheries Center (NWAFC), now the NWFSC and AFSC.

A database management system (DMSII) and the ALGOL programming language were the primary tools used to build the 1981 system. When the system went on-line in October 1981, it included a single type of input transaction that contained data elements, such as date, species, area, gear, port/country/JV, weight-of-catch, number-of-landings, number-of-fish, and dollar-value. The transactions provided by the WDFW, ODFW, and CDFG were data

aggregated on a daily basis, while the transactions provided by the NWAFC were data aggregated weekly.

One very important development during this initial process was the establishment of a set of coastwide PacFIN codes for species, area, gear type, and port/country/JV. Since each data source had its own coding system, it was deemed critical that the PacFIN system be based on a set of codes that would apply throughout the entire geographic range that PacFIN intended to address, as well as across all time periods.

The stipulation that reports present catch estimates for each groundfish species by INPFC area and data source was soon expanded to include the data elements of gear type, port, and month. To meet these additional reporting requirements, it became apparent that the data would need to be summarized as they were received from the data sources and the summaries stored in an on-line "summary" table. This summary-catch table included data for the following: year, time period, species, area, gear type, port/country/JV, pounds, and estimated dollar value. The summary-catch table was essentially a five-dimensional array that allowed for the storage and retrieval of catch and landed value by any combination of time period, species area, gear type, and port/country/JV.

In this summary structure, a time period could represent any month or an entire year. A species code could represent a single species, a species complex, or management group. An area code could represent a single PSMFC area, a single INPFC area, or all areas managed by the PFMC. A gear code could be a single gear type, a group of gear types, or all gear types combined. A port/country/JV code could represent a single port, a group of ports, all ports in a state, or various other combinations of port, country, and JV. The first reports generated from summary-catch tables (PFMC Groundfish by INPFC areas and PFMC Groundfish by Source) were both produced and distributed in October 1981.

Three additional standard reports were developed: Groundfish by Gear Group, Groundfish by Port Group, and Groundfish by Month. All of these initial reports contained coastwide fishery statistics. Because the database contained summaries for nearly every combination of period, species, area, gear type, and port/country/JV, it was decided to enhance the reporting system to produce similar reports specific to each data source. Agency-specific reports were originally intended as feedback to PacFIN agency coordinators so that they would be able to compare PacFIN-compiled summaries with their own agency-generated statistics. However, the state-specific reports quickly became the primary source of landing data for some agency managers, biologists, and economists.

The set of five programs that generated the initial 1981 reports became the primary PacFIN reporting system. The capabilities of this retrieval system continued to be enhanced as new functions and features were suggested by PacFIN clients or were deduced as a result of day-to-day interactions with the various users. Many of the extensions and enhancements that were made in the first few years were a direct result of suggestions and requests made by ODFW personnel.

In January 1982, the GMT requested that the central database include the number of fish tickets classified as groundfish, pink shrimp, etc., so that "indices" of fishing effort could be determined on a coastwide basis. The number of fish tickets, described as deliveries on the PacFIN reports, could be aggregated by each data source, for combinations of management group, area, gear type, and port. Some of the data sources were able to develop the requisite software, but coastwide reporting of delivery information did not commence until March 1987, when all of the PFMC data sources were able to provide data on groundfish deliveries to the central processing system. Nevertheless, delivery statistics by management group became an integral part of the PacFIN system starting in 1982, even though reporting of this information was relegated to agency-specific reports for the first 5 years.

In December 1983, the PFMC began preliminary discussion regarding ways to improve the monitoring component of the PacFIN system, in particular, the timeliness with which landing information was updated and made available to management. The Quota Species Monitoring (QSM) subsystem was largely initiated at the request of the fishing industry, which needed timely information regarding the cumulative catch of the groundfish species regulated by annual quotas (e.g., widow rockfish and sablefish). The industry requested that the cumulative catch estimate be updated monthly, or possibly weekly, to allow them to develop production and marketing plans. At the time, the PacFIN system was providing routine reports on a monthly basis; however, the reports did not contain up-to-date estimates of catch. For example, the most current data provided by an agency were received 15 to 45 days following a

fishing trip, and in some cases data were not received for as long as 4 to 5 months following the actual catch dates.

Initially, the QSM program was administered by one member of the GMT, who was responsible for compiling weekly catch reports following phone conversations with state agency personnel regarding recent catch information for species regulated by quotas. The weekly catch reports for a species were multiplied by a correction factor based on the assumption that the catch data were incomplete. At first, correction factors were determined subjectively, but as more and more catch reports were prepared, improved correction factors became available. Eventually, correction factors were computed by comparing annual summations of weekly reported catches to comparable data in the PacFIN central database. The QSM program was applied using these manual procedures through October 1985. An automated version of the QSM subsystem was finalized in November 1985, which resulted in a more efficient process and more accurate reports than were possible using the original manual operation.

6.2.2 Other Data

The PacFIN system also includes various databases that are not directly associated with groundfish data collected from shoreside landings in Washington, Oregon, and California. In November 1982, the PacFIN system was expanded to include foreign country and JV data for the Alaskan fisheries managed by the North Pacific Fishery Management Council. These data, which were collected and processed by the NWAFC, represented weekly estimates of landed catch by species, area, gear type, and foreign country or JV. Foreign country and JV data are included within the PacFIN system starting in 1981. In April 1984, groundfish data from the ADFG were included in the PacFIN central database. With the addition of the ADFG data, the PacFIN database included catch statistics for all fish harvested from U.S.-controlled waters (0-200 miles) from California to Alaska and landed at U.S. ports.

In July 1986, the Alaska Regional Office (AKR) of the NMFS became a PacFIN data source. The inclusion of AKR data was an important expansion for PacFIN, because these data included catches from domestic at-sea processors that were not being landed at U.S. ports, but were being shipped directly to foreign markets. In recent years, the AKR has primarily submitted data on retained catch for species, or species assemblages, by area, gear type, and week.

In 1987, the PacFIN system also began to receive catch, effort, and economic statistics for the Pacific salmon commercial fisheries of Washington, Oregon, and California. The salmon database, which includes the years 1981 to the present, is a repository for commercial catch statistics associated with the salmon fisheries off the U.S. Pacific

coast (Washington, Oregon, and California). The salmon database includes landings (in pounds, rather than in numbers of fish) and landed value by species, state, port, gear type, and month.

In February 1983, an investigation was conducted to determine if the PSMFC Data Series and the PacFIN system were both necessary. The Data Series consisted of a set of tables (hard copies) dating back to the mid-1950s that included catch statistics for species of groundfish and shrimp. These data were provided by the Department of Fisheries and Oceans (DFO) of Canada, and the four state fishery agencies of the U.S. Pacific coast. Throughout the 1960s and 70s, the Data Series represented the most extensive databases containing coastwide groundfish catch and effort statistics, and these data were the primary information used to construct the broadly utilized Technical Subcommittee (TSC) reports (see section 1.2 for a detailed discussion regarding the Data Series and TSC). It was decided that with certain enhancements to the PacFIN system the Data Series could be eliminated. The following additions to the PacFIN system were deemed necessary: 1) fishing effort (in trawl hours) by PSMFC area and month, 2) logbook-adjusted estimates of catch by PSMFC area and month, 3) logbook-adjusted estimates of species composition, and 4) groundfish catch data from the DFO.

In May 1987, the PacFIN central database was expanded to include fishery-related data collected by the Department of Fisheries and Oceans (DFO) of Canada. The DFO PacFIN coordinator submits catch data twice a year to the PacFIN system. The DFO database within PacFIN currently includes the years 1981 to the present. By 1987, only two of the five fishery agencies were able to provide effort data in trawl hours and logbook-adjusted estimates of catch. The Data Series and PacFIN system merger was never completed in its original form. Although this merger was generally unsuccessful, it did initiate three very important advances in the evolution of the PacFIN system. First, the DFO became a PacFIN data source and user. Second, in 1988, the PacFIN system started providing an annual report to the TSC that included domestic groundfish landings for the entire North American Pacific coast. Finally, the attempted merger of these two coastwide databases generated information that was beneficial during the specification and development stages involved in the redefinition project of the PacFIN system (see below).

6.3 PacFIN System After 1987

6.3.1 Redefinition Project -- Specification

Although the PacFIN system had undergone significant improvements since its inception in 1981, additions and revisions to the central database were needed. First, the ability to provide input data to the central database at the PSMFC-area level was

never fully achieved on a coastwide (including Canada) basis. That is, not all agencies could provide their data in a format that allowed the distributions of catch by area and species to be combined with the fish ticket data, which was necessary to produce the aggregated-catch transactions for input to the PacFIN system. Another shortcoming was the inconsistency between the PacFIN central database and the Research Database associated with the Southwest Regional Office of the NMFS. In 1987, the Research Database was the only coastwide (Washington, Oregon, and California) data system that contained individual fish ticket and vessel data. Fishery researchers and managers occasionally found contradictory information within these two databases, which complicated analyses that utilized these data. A third unresolved issue was that the PacFIN system did not include specific market categories for species of rockfish. Finally, fishery economists had recommended for some time that a primary consideration while developing and expanding the PacFIN system should be the inclusion of all species of fish that are commercially harvested from U.S. Pacific coast waters.

The above requests led to what has generally become known as the "redefinition project" for the PacFIN system. The Data Committee appointed a subcommittee in December 1988 to investigate the feasibility of redefining the PacFIN system, and then to proceed with specification and development stages if the project was deemed a viable one. The subcommittee solicited input from various users of the PacFIN central database regarding their data needs and ways to improve the overall system. The most important requests are summarized below.

1. Groundfish catch statistics by PSMFC area for all relevant data submitted by the participating fishery agencies.

2. Catch, effort, and economic data for all species (not just groundfish) commercially harvested from marine waters off the North American Pacific coast.

3. Fish ticket data on an in-season basis.

4. A historical database that contains detailed fish ticket information that has not been summarized or reduced.

5. Detailed species-composition data that have not been summarized or reduced.

6. A policy whereby agencies submit all fishery-related data to a single, centralized database system (e.g., PacFIN), which would eliminate the need to resolve inconsistencies in multiple databases that reside in various locales.

6.3.2 Redefinition Project -- Development

In October 1990, the Data Committee authorized implementation of the new system, and development began in January 1991. The PacFIN office, in conjunction with the PSMFC, decided to employ a private computer software company to help with the development. By April 1992, the redefined system was able to correctly process all of the 10 new transaction types, and by April 1993, all data for 1981-91 were incorporated into the new, redefined central database.

Following the modifications to the transaction-processing portion of the PacFIN system, the focus turned to developing new software to summarize the fish ticket data and combine the summaries with data on species composition and catch by area. A primary goal of the summarization procedures was the re-creation of the summary-catch tables for 1987-92 (see section 6.4.2 for a description of these data and tables). In the redefined system, the summary-catch tables are built directly from the fish ticket, species composition, and catch by area information submitted to the PacFIN system by the fishery agencies. Summary-catch tables (1987-92) for the U.S. Pacific coast states (Washington, Oregon, and California) became available from the redefined PacFIN database in October 1994.

The first significant users of the new, redefined system was the Groundfish Permit Office of the Northwest Regional Office (NWR) of the NMFS. During 1993, the staff of the NWR verified the fishing history for at least 950 groundfish permit applications using the redefined database. The development process of the redefinition project was formally completed in early 1994.

6.3.3 Vessel Summaries Subsystem

A number of other useful applications based on the redefined system were implemented long before the project was entirely completed. One of these was the "vessel summaries" subsystem. As the redefinition project was being developed, economists involved with the U.S. Pacific coast fisheries, as well as the U.S. Coast Guard, generally recommended that a set of catch summaries by vessel be generated for the years 1981 to the present. The newly requested summaries, which would be expanded later to include detailed information regarding the vessels, were intended to replace the vessel summaries that were originally part of the Research Database located at the Southwest Regional Office.

The subsystem produces and maintains two kinds of files. The vessel summary file contains aggregated landings and landed value information, and 13 other descriptors for each vessel; the trip-principal file contains other characteristics of the fishing vessels, such as principal port, principal gear type, and principal species landed.

Monthly and weekly vessel summaries were being distributed by September 1993. After a few enhancements had been incorporated, based on responses from economists and other interested parties, the vessel summaries project concluded in February 1994. Since then, the subsystem has received only minor changes.

6.3.4 Transition to UNIX/Oracle Computing Environment

In May 1993, the NMFS formally announced that all of the primary computing resources affiliated with the agency would be replaced by homogeneous hardware and software, namely the UNIX operating system and the Oracle relational database management system. This stipulation meant that the overall PacFIN system would need to be restructured, including discontinuing the current Unisys B7900 computer system. The new system, generally referred to as "Orca," was first made available to users, such as PacFIN, in February 1994, but because of various difficulties with configuring the Oracle software, effective use of Orca started in September 1994.

The PacFIN system resided on the same computer system from February 1981 until March 1995, when the Unisys B7900 system was shut down permanently. Prior to this event, the system had been redesigned for the Oracle environment, all necessary tables had been created, and all data had been transferred to the UNIX system and loaded into Oracle tables. Development and testing continued during 1995, and by March 1996, the transition was completed, including all changes that needed to be made to the QSM subsystem, transaction processing systems, aggregated-catch summaries, as well as fish ticket, species composition, and catch by area tables, which collectively are used to produce the final summary-catch tables.

As of August 1996, two major subsystems have yet to be converted to the new Orca system: the vessel summaries subsystem, and a suite of retrieval programs that generate standard reports. Both of these subsystems, once completed, will include capabilities otherwise not available to the PacFIN user community.

6.3.5 Limited-Entry Permit Subsystem

In August 1995, the PacFIN system began accepting limited-entry permit data that were being collected from specific groundfish fisheries. In October 1995, it became possible to access and retrieve limited-entry permit data from the PacFIN system. The limited-entry permit data are collected from each applicant by the Permit Office of the Northwest Regional Office, stored in a computer system developed and maintained by the Permit Office, and then submitted to the PacFIN system twice a month.

6.4 Current PacFIN System

6.4.1 Overall Data Flow

All information included in the PacFIN system is received from the following data sources: four state fishery agencies (ADFG, WDFW, ODFW, and CDFG), two NMFS Regional Offices (AKR and NWR) and a NMFS Science Center (AFSC), U.S. Coast Guard (USCG); and Department of Fisheries and Oceans of Canada (DFO). Information contained in the PacFIN system is originally submitted as a transaction type or as a data file by one of the nine data sources above (Table 6.1).

All data destined for the PacFIN central database are imported into the Orca computer system using one of five methods: 1) file transfer directly into Orca using Internet communications initiated at either the sending or receiving end, 2) file transfer to a computer bulletin board at the PacFIN office and via the Internet to Orca, 3) diskette delivered to the PacFIN office, with the data then transferred to Orca, 4) 8-mm UNIX tape containing an ASCII file, with subsequent data transfer to Orca conducted by operations staff at the PacFIN office, or 5) 9-track tape, which in recent years has not been used by the data sources to transfer data.

Data are submitted or updated at different times, depending on the agency. Data are submitted monthly by WDFW, ODFW, and CDFG. The AFSC submits data on a weekly basis for the Pacific hake vessels that process catches at sea. The ADFG and AKR also provide data weekly. The DFO is scheduled to provide preliminary data each May for the previous calender year, with a final update due in November. Data for the limited-entry history file are submitted twice-monthly by the NWR. The USCG provides the vessel data file annually. It is important to note that data are not always submitted according to the above schedules. For example, data have arrived in the PacFIN office as much as a year behind the agreed-to schedule. For the most part, agencies submit their data in a timely fashion, e.g., the WDFW, ODFW, and CDFG consistently provide data by the 14th of each month. Data completeness varies, however. The ODFW data are normally 90-95% complete 15 days after the end of each month, whereas CDFG data are usually about 90% complete 2-3 months after the end of each month.

6.4.2 PacFIN Database Tables

All of the data submitted by the fishery agencies are validated, to some degree, and then stored in PacFIN database tables (Table 6.2). For purposes of brevity, database tables have been grouped into broad categories, and descriptions are general and primarily applicable to groundfish fishery data that are submitted by the WDFW, ODFW, and CDFG.

Code list tables

Code list tables for descriptors such as species, area, gear type, port/country/JV, and agency are created with data that: 1) originated from within the PacFIN system, or 2) have been submitted by the agencies (data sources). For the most part, code list tables are used to validate, update, retrieve, and provide general descriptions of the data that populate the central database of PacFIN. Currently, there are 11 code list tables: 6 that are created from data originating within the PacFIN system, referred to as species (SP), area (AR), gear type (GR), port/JV/country (PC), agency (AG), and code list (CL) tables; and 5 that address the relationships between codes created by the agencies and codes created within PacFIN, referred to as agency-species (ASP), agency-area (AAR), agency-gear type (AGR), agency-port (APR), and agency-processor (APC) tables.

The species codes used in the various PacFIN database tables, such as the fish ticket tables discussed below, do not necessarily denote a single species of fish, but may refer to a collection of species that have been landed within a single market category. For example, although fish ticket information submitted by the ODFW and CDFG contains a listing for yellowtail rockfish, this reference is actually a market category that contains primarily yellowtail rockfish, but often includes other species as well. Market categories for rockfish are sampled by the state fishery agencies to determine the actual species composition of the categories (see sections 2.4, 3.4, and 4.4 for detailed discussions of the species-composition sampling programs conducted by the individual states). However, in other database tables, species codes do refer to a single species. For example, in the proportion tables described below, the proportion estimates for rockfish are based on additional data that have been collected from the sampling programs for rockfish species composition.

Fish ticket tables

Fish ticket information provided by the fishery agencies is included primarily in two database tables. The fish ticket (FT) table contains delivery-specific information, where each row of the table contains attributes of a completed fish delivery. The fish ticket lines (FTL) table contains market category-specific information, where each row of the table contains attributes of the market categories included on a corresponding fish ticket.

Proportion tables

The PacFIN central database contains three tables that are collectively referred to as proportion tables (ACM, SCM, and ECM tables). The data contained in these tables, along with data from the fish ticket tables, largely distinguish the redefined PacFIN system from the earlier system.

The catch-by-area composition (ACM) table contains proportions that are used to distribute catch to PSMFC areas for specified "strata" (e.g., species/port/gear type/time period). A fifth attribute, grade (size of fish), is commonly included with the four attributes above to specify strata for landings of sablefish. Catch-by-area transactions are used by the WDFW and ODFW, but are not currently utilized by the CDFG, which uses ports to identify specific PSMFC areas where catches were made (see sections 2.6, 3.6.1, and 4.6.1 for discussions regarding procedures used by the individual states to apportion catches to geographical areas).

The species-composition (SCM) table contains proportions that are used to distribute catch to individual rockfish species for specified strata (e.g., rockfish market category/port/gear type/PSMFC area). Species-composition transactions are used by the WDFW, ODFW, and CDFG. The estimated proportions of species composition for rockfish market categories are determined from data collected through sampling programs conducted by the individual states (see sections 2.4, 3.4, and 4.4 for further discussion regarding these data collection programs).

The effort-by-area composition (ECM) table contains proportions that are used to distribute effort to PSMFC areas for specified strata (e.g., management group/port/gear type/time period). Management group refers to the actual fishery that the data were collected from, such as the groundfish, salmon, or shrimp fisheries. This table is similar to the catch-by-area composition table, but the fishery descriptor that is being apportioned is effort (number of deliveries and trawl hours) rather than catch. Effort-by-area transactions are currently used only by the ODFW.

Summary tables

Catch and effort statistics that have been summarized within the PacFIN system are included in four primary database tables: summary-catch (SC) table, detail-catch (DC) table, summary-effort (SE) table, and detail-effort (DE) table. All statistics contained in these tables are derived from detailed information residing in other areas of the PacFIN central database.

The summary-catch tables contain reduced catch statistics that can be retrieved easily and quickly. The detail-catch tables are similar and related to the summary-catch tables. During 1981-86, aggregated-catch transactions (as opposed to individual fish tickets) were the only types of transmissions that could be used to submit catch data to the PacFIN system. The ADFG, AKR, AFSC, and DFO continue to use this process to submit catch data. Currently, the WDFW, ODFW, and CDFG use fish ticket and proportion tables to submit catch data, which are then modified into aggregated-catch transactions internally within the PacFIN system. The data contained in aggregated-catch input records are permanently stored in the detail-catch tables. The detail-catch

tables (1981-86) for the WDFW, ODFW, and CDFG contain the original daily aggregated-catch records submitted by the respective agencies. The detail-catch tables (1987 to the present) for the WDFW, ODFW, and CDFG contain internally generated monthly aggregated-catch statistics. The summary-catch and detail-catch tables are consistent with each other for all years 1981 through the present.

The summary-effort (SE) and detail-effort (DE) tables contain three "measures" of fishing effort: number of deliveries, trawl hours, and days fished. Number of deliveries is essentially the number of fish tickets. Trawl hours is an estimate of the number of hours a fishing vessel is actually engaged in the act of fishing with its net in the water. Days fished is derived from information included in the fish ticket tables.

The summary-effort and detail-effort tables are structured in a similar fashion as the summary-catch and detail-catch tables, with the important distinction that summary statistics for effort can be obtained for certain management groups, but not for individual species. For example, the SC and DC database tables contain summarized catch statistics for sablefish, but do not include information regarding the number of deliveries or trawl hours associated with sablefish catches. The effort data for sablefish are combined with effort statistics for other groundfish species and presented as a single value for the entire groundfish management group, which is included in the SE and DE database tables. Summary-effort and detail-effort database tables are available for the years 1981-94; however, because these tables are currently receiving modifications, 1995-96 data are not yet available.

Other tables

The PacFIN central database includes several tables that are used in conjunction with the tables described above to produce various statistics on a routine or requested basis. Statistics generated through the Quota Species Monitoring and limited-entry permit subsystems are contained in QSM and limited-entry permit tables, respectively (see sections 6.3.5 and 6.4.5 for further discussion regarding these two subsystems within PacFIN). The state-vessels (SV) database table contains information regarding the commercial fishing vessels registered by each state to harvest fish. The USCG vessels (CG) database table contains selected attributes from the USCG's Merchant Vessels data file. Some of the vessel attributes included in this table are gross weight, length, horsepower, and the year the vessel was built. The non-vessel (NV) database table is an ancillary table that contains vessel identification information, which is created when the SV table is used to translate agency vessel plate numbers to either a USCG vessel identification number or a state marine board identification number.

The average-weights (AW) database table contains estimates of average weight that are subsequently used to calculate total estimates of the number of fish landed within a specified strata, such as species/port/gear/PSMFC area. The average-weights table is used exclusively by the ODFW for species of salmon, sturgeon, and shad.

The update-log (UL) and detail-log (DL) database tables contain information that is generated during processing operations within the PacFIN system. The dates associated with amended database tables are stored in the update-log table. The amount of data that enters the system following an update is included in the detail-log table; this information is subsequently used to determine how complete data transmittals are for each PacFIN data source (see section 6.4.5 for further discussion regarding data completeness).

6.4.3 Central Processing -- Update

Update processing within the PacFIN central database is now conducted within a UNIX/Oracle environment. The suite of "update" software utilized in the PacFIN system is composed of the following programs and languages: Oracle's PL/SQL, Oracle's SQL*Plus, Oracle's SQL*Loader, and the 'C' programming language. Data submitted by the agencies are validated,

to some degree, during update processing. Agency transactions that are "flagged" as invalid are reviewed by the agency's PacFIN coordinator, who is responsible for resolving the errors.

Although the central processing system includes some routines to validate submitted data, the content of each data file (i.e., the value of each datum) is strictly the responsibility of the individual agencies. That is, although input data are generally reviewed for possible errors, the central processing system does not include comprehensive validation routines at this time.

For the original fish ticket lines (FTL) data provided by WDFW, ODFW, and CDFG, the redefined PacFIN system includes update routines that provide estimates of landed value for catches that do not include price information. The FTL rows containing actual prices are used to build a temporary table of information on total pounds and landed value classified by market category, condition (e.g., dressed vs. whole), disposition (e.g., animal vs. human food), grade, port/country/JV, and gear type. This table of actual prices is then searched to determine an estimated price for each FTL row that is missing a price. Similar procedures are used to derive estimated landed values from the aggregated-catch transactions provided by ADFG, AKR, and AFSC. The DFO data source does not provide any economic data, and the PacFIN system does not attempt to estimate the landed values of Canadian catch transactions.

An important focus of the redefinition project was to improve and streamline methods for generating aggregated-catch statistics, which inherently involved modifications to internal summarization procedures for fish ticket data. The catch-by-area and species-composition data that are received by the WDFW, ODFW, and CDFG are applied to summarized fish ticket information (FTL data) to provide aggregated-catch statistics on a monthly basis. The following steps are used to process catch data that are received by the agencies: 1) monthly aggregates of FTL data are created, 2) catch-by-area proportions (ACM tables from agencies) are applied to the aggregated-FTL data, 3) species-composition proportions (SCM tables from agencies) are applied to both the aggregated-FTL data and summations resulting from the application of catch-by-area proportions to certain FTL aggregates, and 4) aggregated-catch transactions are generated.

The computation of monthly aggregates of FTL data begins by identifying those months that need to be summarized. This is determined by finding each month that occurs at least once in the set of FTL, ACM, and SCM transactions that have been processed. The aggregated fish ticket data are summarized by month, species, port, gear, and PSMFC area. This aggregation contains the following attributes of the catches: round-weight equivalent pounds, number of landings, number of fish, pounds that were actually priced, and estimated landed value. Statistics for the number of landings and the number of fish may not be available for certain species/species groups. It should be noted that all FTL data for the selected months are involved in this aggregation exercise, not just those that will be subsequently apportioned.

Catch-by-area (ACM) proportions that are submitted by the agencies are then applied to the monthly aggregations of fish ticket data for only those cells (i.e., month/species/port/gear type/PSMFC area) that have corresponding proportions in the ACM table. Note that many of the monthly aggregations of fish ticket data do not generally need to be apportioned by the catch-by-area proportions. The catch-by-area (ACM) proportions submitted by the agencies are then applied to the monthly aggregations of fish ticket data, but only for those aggregate combinations that have corresponding proportions in the ACM table. In general, many of the monthly aggregates do not need to be apportioned to catch by area because the data for area of capture on the fish tickets cannot be further refined. Those monthly aggregates that need adjustment for catch by area are aggregated over area, and the ACM proportions are applied so that each month/species/port/gear type aggregate is partitioned into one or more month/species/port/gear type/area aggregates. Currently this process only apportions the data into PSMFC areas. The new aggregates replace the original ones so there is no "double counting" and no changes in the total pounds landed. Data on number of landings are set to null for any aggregation derived by applying either catch-by-area or species-composition proportions. The present system makes

adjustments for catch-by-area to three categories of groundfish data: WDFW data from Puget Sound, WDFW from coastal waters, and ODFW data with area equal to "unknown."

At the next step of processing, the monthly aggregates for rockfish market categories are apportioned into monthly aggregates by rockfish species. This processing is applied both to the aggregates that were apportioned to area and those that were not. Species-composition proportions (in table SCM) are currently applied only to rockfish market categories, but similar proportions could be applied to any market categories for which the agencies were able to provide species-composition proportions. The SCM proportions are applied by matching them with the monthly aggregates of fish ticket data based either on month/species/gear type/area/port or based on month/species/gear type/port. The aggregate values for round weight, number of fish, pounds priced, and estimated landed value are apportioned by multiplying each by the corresponding SCM proportion. Data on number of landings are unavailable for any aggregations that were derived by applying either ACM or SCM proportions.

The fourth and last step of update processing generates aggregated-catch transactions to update the summary tables (DC, DE, SC, and SE); for example, to delete outdated summary statistics and replace them with recent statistics. The sources of these transactions can either be the agencies (ADFG, AKR, AFSC, or DFO) that directly provide aggregated data or the previous steps in the summarization process (for the FTL data provided by WDFW, ODFW, and CDFG).

The following example illustrates the process. Suppose new landing data are received. The new information is validated and then inserted into the detail-catch (DC) and detail-effort (DE) tables. In addition, a copy of each transaction is saved and used to update corresponding summary statistics in the summary-catch (SC) and summary-effort (SE) tables. For each transaction to the DC and DE tables, five vectors are developed to update the data in the SC and SE tables, one each for period, species, area, gear type, and port. For all possible combinations of the items in these vectors, values are generated for the summary-statistics round weight in pounds, number of landings, number of fish, estimated landed value, and pounds with prices. The generated values are then used to modify the information in the SC and SE tables. A single change to the data in

the DC table results in a multitude of changes to corresponding data in the SC table because the information is contained in a suite of alternative summarizations in the SC table.

For example, if the new data represented catches of Dover sole taken in May from PSMFC area 2C by longline gear and landed at the port of Astoria, then the period

vector would include items for the month of May and for the annual period; the species vector would include items for Dover sole, flatfish, and groundfish; the area vector would include items for PSMFC area 2C, INPFC area Columbia, Pacific Council region, and all regions; the gear-type vector would include items for longline, hook and line, and all gear types; and the port vector would include items for Astoria, the Columbia River (Oregon) port group, all Oregon ports, all domestic ports and at-sea processors, and all ports/joint ventures/foreign countries.

6.4.4 Central Processing -- Retrieval

There are two primary methods for retrieving data from the central PacFIN database: using SQL*Plus routines or using specialized reporting programs. SQL*Plus is a general-purpose database query language that is an integral part of the Oracle relational database management system. PacFIN users who have access to the Orca computer system in Seattle can develop their own SQL*Plus routines. Alternatively, the PacFIN staff have developed a large suite of SQL*Plus routines for retrieving information from the PacFIN central database. A document entitled "Using Unix and Oracle to Access PacFIN Data," which is available upon request from the PacFIN office, gives an introduction to these SQL*Plus routines, as well as other information for new users. These routines can be used to retrieve selected data or can be used as templates for users who wish to develop their own custom retrievals.

The other mechanism for retrieving information from the PacFIN database is to use one of the six reporting programs that have been developed as exact replacements for the reporting programs that were part of the earlier PacFIN system. Examples of the reports produced by these programs, which have become known as the "PacFIN standard reports" (see section 6.2.1), can be found on the PSMFC homepage on the World Wide Web (http://www.psmfc.org/). As of this writing, the subsystem for generating PacFIN standard reports is still in development. When it is complete, Orca users will be able to generate their own standard reports; but until then, selected standard reports will be produced by the PacFIN office and made available as described above (see section 6.2.1).

6.4.5 Data Completeness

Data completeness for each PacFIN data source is determined using a variety of indicators; two straightforward methods are presented here. One method involves tracking the amount of data that enters or leaves the PacFIN system during update processes. The detail-log database table includes the total pounds that have been added (or deleted) for each month for all groundfish transactions. Another method used to help determine data completeness is to compare the historical catches that are presented in the summary-catch tables. For example, the monthly totals for catch for

the most recent year can be compared to catches from earlier combinations of year and month to obtain rough percentage estimates of completeness in the most recent year.

6.4.6 Confidentiality of Data

The PacFIN central database contains "confidential" information, where the economic history of individual fishing vessels and fish processors can be determined from the contents of the fish ticket tables (FT and FTL tables).

Access to confidential data is regulated through rules established by NMFS, under the National Oceanic and Atmospheric Administration (NOAA). The rules stipulate that the only information that can be made available to the general public are those statistics that do not reveal the economic activity of individuals or corporations. In order to adhere to the confidentiality rules set forth by NOAA, the PacFIN office requires users of confidential data to sign a "Certificate of Non-disclosure of Confidential Fisheries Data." Access to confidential data is restricted to individuals participating in PFMC activities that require the use of confidential information. However, other individuals who have contracted with the PFMC on particular projects are also granted access to confidential data, given they sign the above Certificate and agree to destroy the data after completing the study. Only employees of NMFS and other Data Committee member agencies are considered for on-line access to the PacFIN system.

6.5 Acknowledgments

This chapter was prepared with the assistance of Ed Heyman, Mana Hung, and Brad Stenberg. In addition, Stan Allen was instrumental in the completion of this chapter. We especially thank Drs. John Harville and Harvey Hutchings, who as the PSMFC executive director and the NMFS/NWR Fisheries Management Division chief, respectively, were two of the driving forces behind the PacFIN concept and project during the early formative years from 1978 through 1983; the success of PacFIN is a result of their vision and tenacity.

6.6 Citation

Daspit, W. P. 1996. A description of the Pacific Fishery Information Network (PacFIN) 1981-1996. Unpubl. manuscr., 39 p. plus appendices. (Available from Pacific States Marine Fisheries Commission-PacFIN F/AKC, Building 4 - Room 2066, 7600 Sand Point Way NE, Seattle, WA 98115.)

Table of Contents