SUMMARY OF WEST COAST GROUNDFISH OBSERVER PROGRAM DATA COLLECTION AND QUALITY CONTROL PROCESS NORTHWEST FISHERIES SCIENCE CENTER

Summary

The Northwest Fisheries Science Center's West Coast Groundfish Observer Program regularly provides the Pacific Fishery Management Council with discard and bycatch rate information. The rates are determined from data collected by at-sea observers deployed on commercial fishing vessels operating in west coast fisheries catch and retain groundfish. Before use, observer data passes through four phases, which take a total of three and a half to four months. This document details these processes, their timelines and suggests possibilities for reducing the amount of time required to complete them and summarizes difficulties in moving toward in-season catch accounting.

The first phase, data collection, entry and initial quality control, takes two months from the collection of the last observed trip of a period. This time could be reduced somewhat if fewer observers were deployed and instead resources were used for staff to quality control data. Since this would reduce existing coverage levels of groundfish fisheries this is not considered an option

The second phase, identifying and attaching corresponding fish ticket with observer data, takes about 10 weeks from the last observed trip of a period. Currently, the program waits two months after the last observed trip of a period for ~90% fish ticket completion rate in the PacFIN system to begin locating and attaching fish ticket and observer data. The time required for this phase could be shortened if fish ticket data were available more rapidly.

The third phase, data processing and analysis, takes about four weeks for each fishery for which discard and bycatch rates are produced. The time needed to complete this phase could be reduced if the fish ticket data set was entirely complete in PacFIN data system sooner, each landing could be easily attributed to a specific fishery, and there were greater consistency in recording catch categories in fish ticket and logbook records. Production of reports that focus solely on summarizing new data, rather than also evaluating changes from previous reports, would also expedite this phase.

Within the fourth phase, final preparation of the data and updating of projection models for use in management, takes about two weeks for the trawl fleet and an additional week for each of the other fleets. Estimation of total discard mortality for each fleet during the previous year requires an additional two weeks for the trawl fleet and another week for each of the other fleets.

In summary, the principal objective of the observer program is the documentation of discards and is not to duplicate the other existing data systems that estimate the catch that is retained. Accordingly, observer data represent only a portion of the information required to estimate total fishing mortality. Discard information must be combined with data from other sources such as fish tickets and vessel logbooks (where available), in order to document total catch. While some internal steps can be taken to reduce the data turnaround time in releasing observer data, because

of its connection with other data sets, other programs and agencies must also take steps in order to reduce the turnaround time.

There are many issues that complicate the potential for processing observer data on a more real-time basis and using those data for in-season management of fleet or vessel total catch caps. In addition to the availability of auxiliary data, substantial concerns about how data would be applied to unobserved vessels would need to be addressed. Existing methods for pooling observer data to protect confidentiality and achieve reasonable sample sizes would not function well in real-time mode, given existing levels of observer coverage. Changes in the sampling design used to distribute observer coverage might improve the extent to which all fleet fishing strategies are observed within any particular 2-month period. However, further analysis is required in order to evaluate the degree to which gaps in coverage would create problems for reporting and applying rates from observed to unobserved vessels over short time periods throughout the year.

Overview

The Pacific Fishery Management Council manages the fisheries off the coasts of California, Oregon and Washington. To aid in Council decisions, the Northwest Fisheries Science Center (NWFSC) provides fisheries research and data on species populations, bycatch rates and economic impacts. The West Coast Groundfish Observer Program (WCGOP) is responsible for collecting the main source of groundfish bycatch data provided to Council. This document describes how the bycatch data is produced for management use. The multifaceted process and the steps involved in final production of data are described and suggestions are made on actions that could be taken to reduce the time to produce the final bycatch and discard information.

The WCGOP deploys trained fishery observers in the limited entry bottom trawl, limited entry fixed gear and the 'open access' fleets (which currently includes the California sablefish open access fleet, Oregon and California near-shore fleets and a portion of the California halibut trawl fishery). Observers' primary responsibility is to monitor the composition and volume of discarded catch on each tow/set. The program randomly selects portions of each fleet for coverage based on geographical area, fishery, vessel activity, and the availability of observer resources.

Vessels participating in the limited entry trawl, limited entry zero tier fixed gear, near-shore and/or California halibut fisheries are selected for observer coverage for all trips taken during a two-month trip limit period. As the limited entry fleets land the majority of groundfish, they are a higher priority and typically all limited entry trawl and zero tier fixed gear trips are observed during their selected period. Coverage of the near-shore and California halibut fisheries are given at this time a lower priority, therefore observers cover these vessels only if a limited entry vessel is not active. Vessels participating in the limited entry sablefish endorsed fishery are selected for all trips during the sablefish season where the vessel lands sablefish against a tiered quota. Typically, all trips with sablefish tiered landings are observed.

The following section highlights the major steps or phases the data goes through before it is submitted for use by the Council. Figures 1 and 2 provide a visual illustration of these phases as well. Figure 3 provides an example timeline of the entire process for one fishery.

Phase 1 - Data Collection, Entry and Initial Quality Control

Observers collect fishing effort information, estimate total catch weight, estimate discard weight, and take species composition and biological samples from the discard (sampling details are available from the WCGOP Observer Training Manual at: http://www.nwfsc.noaa.gov/research/divisions/fram/observer/index.cfm). This information is recorded on waterproof deck forms and upon completion of a trip; the observer enters the data into the observer database. Once an entire trip is entered, the observer runs a data integrity check program consisting of approximately 170 data error checks. At the end of each month, the observer sends the hardcopies of their data to their debriefer. The debriefer, a staff member, then reviews sampling methods and checks for any calculation or data form errors. The data is returned to the observer who corrects errors on the hardcopies and in the database. After a two-month period is completed and a debriefer has checked all data, the observer is interviewed. During the interview, a last check is done for any more data or sampling method errors, species

identifications are verified, all biological samples are logged and verified, and the observer's performance is evaluated. After the interview, the debriefer evaluates the data quality, either passing or failing a trip or a segment of trip based upon standard evaluation criteria. The debriefer performs a final verification to confirm that the hard copies of the data match the data in the computer. Finally, the debriefer runs a data error check for the data set, updates any errors found, and closes the trip to prevent further updates.

Time needed:
It typically requires two-months from the end of the data collection period to complete this phase.

Suggestions for reducing the time needed for phase 1:
Currently, the observer program does not release any preliminary data. As the data is from a sub-sample of the fishery, the program must ensure the highest data quality. Accelerating the schedule may impact data quality and the ability to maximize sea time as observers may have to forego trips for debriefing.

Phase 2: Identify and Attach Corresponding Fish Ticket Data with Observer Data

Once phase one is completed, the database manager runs a final quality control check on the entire data set to be analyzed. The quality control can be done by fishery or as a combination of all data collected from all fisheries during a given time period. A debriefer reviews all data that was flagged from the data error check and makes updates when necessary.

After the quality control, the observer data is matched with fish tickets (downloaded from PacFIN) on a per trip basis. To estimate total catch on observed vessels and to relate the data to the unobserved portion of the fleet, observer data and fish ticket(s) data are matched (i.e. the fish ticket(s) data produced for every observed trip is collected from PacFIN). Although locating fish tickets for all observed trips seems a relatively straightforward process, it is not. The state agencies are responsible for delivering the fish ticket information to PacFIN and there is a delay between fish ticket production at the time of landing and when that information is available for observer program use in PacFIN. Currently, it takes approximately two months to have a 90% fish ticket completion rate.

To aid in identifying the correct ticket data, observers record fish tickets number(s) for all trips observed. Using this information, most observed trips can be easily matched with the correct fish tickets. However, observers often only have one fish ticket number from trips that had multiple fish tickets generated as additional fish tickets can be created due to weighbacks, etc.

The database manager searches through the fish ticket list by vessel and using the landing date recorded by the observer, ensures that all fish tickets generated from a trip are located. Cancelled tickets are another complicating factor. For example, there are some ports where a fish ticket is generated at the dock but then the fish are taken by van to a processing facility or market. It appears that sometimes when the transportation of fish occurs, the processor completes a second fish ticket, making the first fish ticket obsolete. Finding the correct tickets can be especially difficult, as there are many days in between vessel landing and the fish ticket being generated. The database manager has created a query to make this process more efficient, however it still

takes two weeks to complete this second phase and phase two can not begin until at least a significant portion of the fish tickets from the trip limit periods of the data set are available in PacFIN. Currently, the observer program waits for ~90% fish ticket completion rate or about two months.

Time needed:
Two months and two weeks.

Suggestions for reducing time needed for phase 2:
If the time to submit tickets were reduced, then the time needed to accomplish this phase would also be reduced. Electronic fish tickets are being investigated by at least one state as a viable option to decrease upload time to the PacFIN data system.

Phase 3: Data Processing and Analysis

After the data have gone through quality control, error checking and all fish tickets are located for observed trips, the data are processed by the analyst and then used to estimate discard rates. As fish ticket data are collected on a trip basis and observer data are collected on a set/haul basis, the data sets must be processed to relate to each other.

First, the analyst estimates haul-level weights of species or species groups. As the observer often subsamples hauls/sets, the analyst uses the weights of the species in the each subsample to estimate the weight of the species in the haul/set. For details on methodology, please see the observer data reports available online at:
http://www.nwfsc.noaa.gov/research/divisions/fram/observer/index.cfm).

As previously mentioned, the observer data set and the fish ticket data set use different coding and have different levels of detail. Before the data sets can be joined, the analyst must assign consistent species, catch categories, port and port group coding to the two datasets.

Once the two data sets have consistent coding, the next step is to match catch categories on the fish tickets to retained catch categories in the observer data. The catch categories are species or species groups. The first part of this step involves calculating the trip-level weights of retained for each catch category for both observer data and fish ticket data. The fish ticket and observer data are then matched by catch category and trip. Next the two datasets are joined and the values from the observer data and fish ticket data are compared. At this point, discrepancies between the observer data and the fish ticket data will become obvious.

The analyst works to identify and resolve gross discrepancies between the observer data and the fish ticket data, such as extreme differences in total retained catch and missing fish tickets. The data are also examined for extreme differences in the weight of the landed catch between fish ticket and observer data. Another issue is that catch categories will appear in only one of the datasets; this discrepancy often occurs when catch categories are combined (i.e. other flatfish) in one dataset and specified in the other dataset (i.e. other flatfish vs. butter sole, sand sole, and curlfin sole). For each one of these discrepancies, an attempt is made to determine the cause of the discrepancy to resolve it.

For the gross discrepancies, the analyst attempts to determine the cause. Investigating the discrepancies may involve searching for additional fish tickets for a trip and looking at comments by haul in the observer data. Finally, decisions are made on the course of action for each discrepancy that does not have a clear cause. This matching process is very labor intensive and time consuming, as each discrepancy must be handled individually.

After the causes of the discrepancies are corrected, the two datasets are then joined again. Another check for discrepancies is repeated and if necessary a third or fourth check is conducted.

Once the observer data and fish ticket data are satisfactorily joined by category and trip, the weight of the observer retained catch is adjusted so that the total weight by trip for each catch category in the observer data matches the weights in the fish ticket data. Also for the catch categories only in the fish ticket data, the weight of the catch category landed in each haul/set is estimated. At this point, the data are ready to be analyzed.

For the observer program data reports, the analyst estimates coverage, total discarded catch for observed trips, and discard and bycatch rates for observed trips. To calculate the coverage of a fishery by weight, the total landings made by the fishery are determined by adding together fish ticket weights associated with the fishery. The PacFIN system is queried for this data, but determining total landings for each fishery can be problematic, as the fish ticket does not identify the fishery in which a vessel was participating. Associating landings with limited entry trips is not too problematic for the limited entry fleets as these vessels fish almost exclusively in only one of the limited entry fisheries. In contrast, the open access fishers often participate in many different fisheries, sometimes in the same trip. For the open access fisheries, the program has yet to identify a method to estimate total cumulative landing by fishery. WCGOP is currently working with the states to attempt to estimate the landings by fishery. The inability to estimate total landings for a fishery results in the program being unable to estimate the percent coverage by weight or determine how well observer data matches normal fleet spatial and temporal activity. Also, discard and bycatch rates can be determined for observed trips but any extrapolation to a fishery as a whole can be challenging and utilize multiple assumptions.

For the most part, estimating total discarded catch and discard and bycatch rates of observed trips for each species/species complex is relatively straightforward. Total discard for over 30 species is determined for multiple strata (see Table 4 from observer program trawl report, Sept 2005 at: http://www.nwfsc.noaa.gov/research/divisions/fram/observer/datareport/trawl/datareportsep2005.cfm) by using a ratio estimator to relate the discard and bycatch to hours of towing and/or total retained catch. However, the process is lengthy due to the sheer number of species and strata. A report typically contains 60-80 tables.

Finally, the data and tables are summarized and the reports are written. Due to the fishery-specific differences, the data processing, analysis and writing has to be done separately. Currently four, fishery-specific reports are produced by the observer program.

Time needed:
Four weeks per fishery.

Suggestions for reducing time needed for phase 3:
A complete set of fish tickets would eliminate the need for the analyst to search for missing tickets. A mechanism to denote on a fish ticket which fishery each fish is being landed in would make it easier to determine total landings in a fishery. In particular an identifier added to fish tickets and logbooks to identify when EFP fishing was occurring would be helpful because bycatch rates observed in the general fishery should not be applied to fishing experimental fishing where either retention is required or discards are documented separately. There should be consistency in recording catch category labels on fish tickets between states/processors. This would help data turnaround since the larger the number of differences on how a species is recorded on a fish ticket, the more discrepancies during the matching process that the analyst must investigate.

Phase 4: Development and delivery of discard data and modeling for use in management

Observer data are used in management in two ways: for estimating past amounts of fleet discard, and for projecting discard and bycatch amounts expected in future periods.

For the groundfish trawl fleet, discard in the previous year is estimated using vessel-level data from fish tickets and logbooks, along with observer data from that year. Logbook and observer data are stratified by depth, area, and season. The amount and distribution of fleet effort, obtained from logbook data, is combined with observed discard rates to estimate overall amounts of discard for trips included in logbook records. Because logbooks do not capture 100% of groundfish trawl trips, ratios of fish ticket-to-logbook retained catches are used to expand the discard estimates to the entire trawl fleet. Since non-trawl fisheries lack comparable logbook programs, estimation of past discard relies primarily upon observer and fish-ticket data. Observer data have been incorporated in a depth-specific manner in the sablefish fishery, based on closed-area boundaries, and in the near-shore fisheries, based on state estimates of the distribution of fishing effort. Computation of the best estimates of past discard requires that all of these components be available before the analyses can be conducted.

It must be stressed that discard estimates from WCGOP-observed fisheries represent only a portion of the information needed to estimate total fishing mortality. Fish tickets are needed to document retained groundfish catch in commercial and tribal fisheries. Data from recreational fisheries are needed to document both retained catch and discard mortality in those fisheries. Additionally, estimates of groundfish discard mortality from unobserved fisheries must also be accounted for. Estimates of mortality from this last category, as well as catches attributable to research fishing and exempted fishing permits are typically compiled by the GMT.

The Council is most commonly presented with <u>projections</u> of total catch and discard, as part of GMT evaluation of alternative management measures. The model employed by the GMT (and developed by the NWFSC) to project total catch for the trawl fleet uses 4-year, weighted averages of fish-ticket, logbook, and observer data, where the most recent data are assigned the greatest weight. The data are stratified into depth zones and areas that are consistent with regional and closed-area options commonly considered for Council management action. Projection models for fixed-gear sablefish and near-shore fisheries utilize expected regional landed catches of target species in those fisheries (based on OY allocations or other harvest targets) and prior observer data.

Current time requirements for preparing updated data for projection models:
2 weeks for trawl; 1 week each for other observed fleets.

Current time requirements for developing estimates of the previous year's discard:
2 weeks for trawl; 1 week each for other observed fleets.

Impediments to in-season catch accounting and suggested next steps to be taken by WCGOP to improve data turn around.

Moving towards a more real-time accounting of total catch, whether at the fleet or vessel level, is complicated by a number of issues. With the existing level of observer coverage, discard data are currently pooled across depths, seasons, and large areas in order to ensure that discard rates applied to unobserved vessels are based on reasonable sample sizes. Because the distribution of observers is achieved through a stratified random sample based on only on gear, port group, and 2-month period, and the actual distribution of fishing effort is uncertain, there is no guarantee that observed vessels will fish within specific depth zones or shorter time intervals in which unobserved vessels fish. To protect the confidentiality of observed vessels, current national policy does not even permit discard rates to be reported for strata in which fewer than 3 vessels have been observed. The current schedule for evaluating fleet discard permits observer data to be stratified in a manner that accounts for reporting and statistical criteria after the actual temporal and spatial distribution of observed fishing effort during a year is known.

Applying observed data to the unobserved fleet on an ongoing basis throughout the year, for the most recent week, month, or even 2-month period, would not permit this sort of data-pooling. Aggregating observer data on a coast-wide basis, or across all depths, in order to achieve adequate sample sizes is not an acceptable option, as existing data clearly indicate vastly different rates of encounters and discard for most species across these dimensions. Over some duration of time—likely at least 2 months—gaps in observer data could be reduced, through additional stratification of observer sampling effort, in order to increase the likelihood of collecting data from the full cross-section of fishing activities in which the fleet is likely to engage during that period. But such an effort would need to be based either on prior vessel participation or on an advance declaration process. Neither option would carry a guarantee that an adequate number of observations would be made within a stratum, and the latter would require vessels to commit, at least 4-6 months in advance of a particular period, to conduct at least some of whatever fishing they did within a specified depth stratum. These approaches also imply that vessels participating in strata with fewer total participants are likely to be observed more frequently than is the case with the current sampling strategy.

Without much higher levels of observer coverage, shorter periods of catch accounting would be associated with increasing likelihood that an inadequate number of observations would occur to permit a reasonable analytical stratification of fleet effort to be employed. **New analysis of existing observer and logbook data could be conducted to reveal the likelihood that, with existing levels of coverage, gaps in observer data for specific depth and area stratifications, over a range of time periods, would prevent reasonable application of discard rates to unobserved vessels.** Protocols would also need to be evaluated and agreed upon for revising estimates of discard, in-season, based on pooling accumulated observer data.

The program can also investigate the time taken by observers to submit their data. Currently, observers submit their data within five days of end of a trip. The program can test whether this time can be shortened without missing sea days or overburdening observers. In addition, the program can investigate how many mistakes are made in an observer's original submission of data before data quality control steps. This would give an indication as to the frequency and type of mistakes observers may be making and the steps that would have to be taken to reduce them.

In addition to these types of changes within the program for collecting observer data, fish ticket, and preferably also logbook, data would need to be available on the same schedule as the observer data. Electronic fish tickets would facilitate an accelerated ability to document the landings of observed and unobserved vessels. However, without equally rapid documentation of fishing depths, the estimation of discard would require the development of new models dependent upon algorithms for estimating the depth of fishing based on past participation and/or species composition. Use of an algorithm in place of logbook information could lead to larger error in the initial bycatch estimates, and increase the need for re-estimation once logbook data become available.

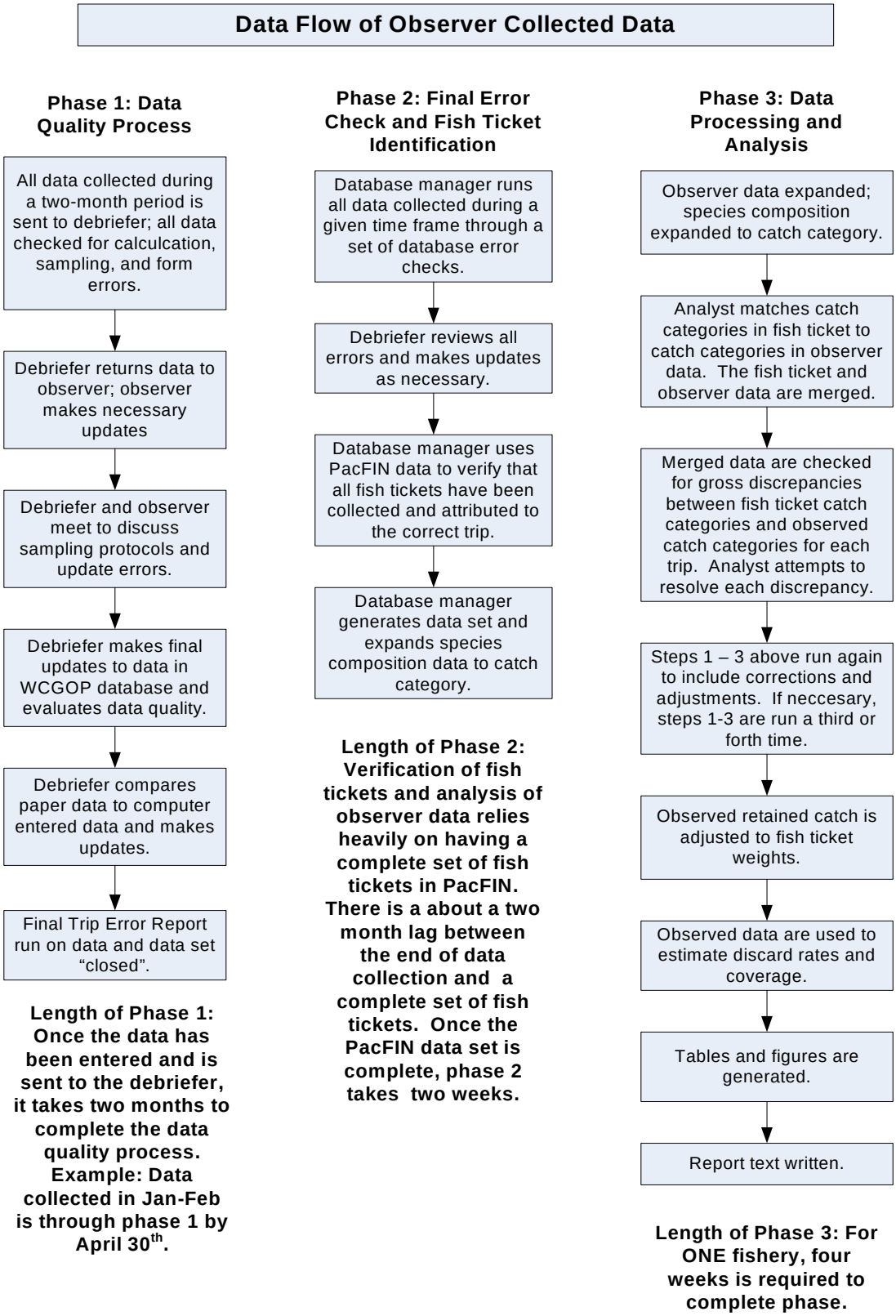

Figure 1. Overview of observer data collection, processing and analysis

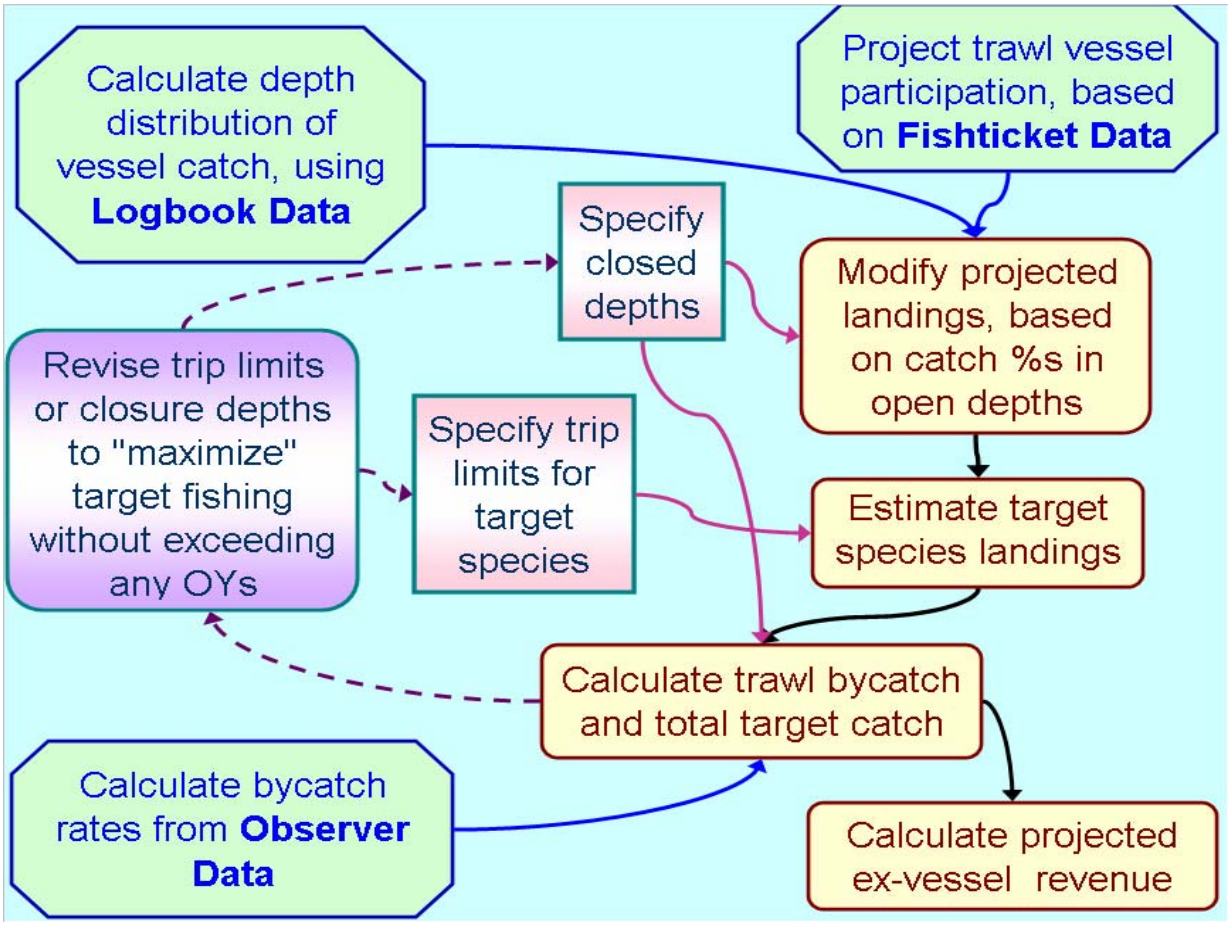

Figure 2. Schematic of the trawl bycatch model (phase 4) used by the Groundfish Management Team (model varies based on available data for each fishery)

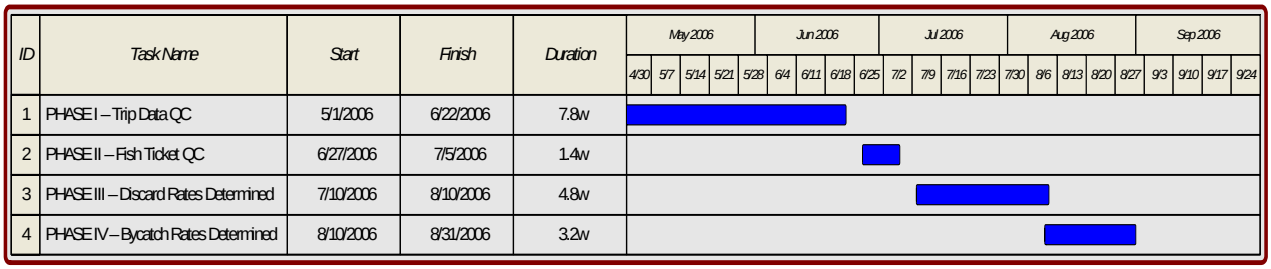

Figure 3. Example timeline for the completion of bycatch analysis for one fishery