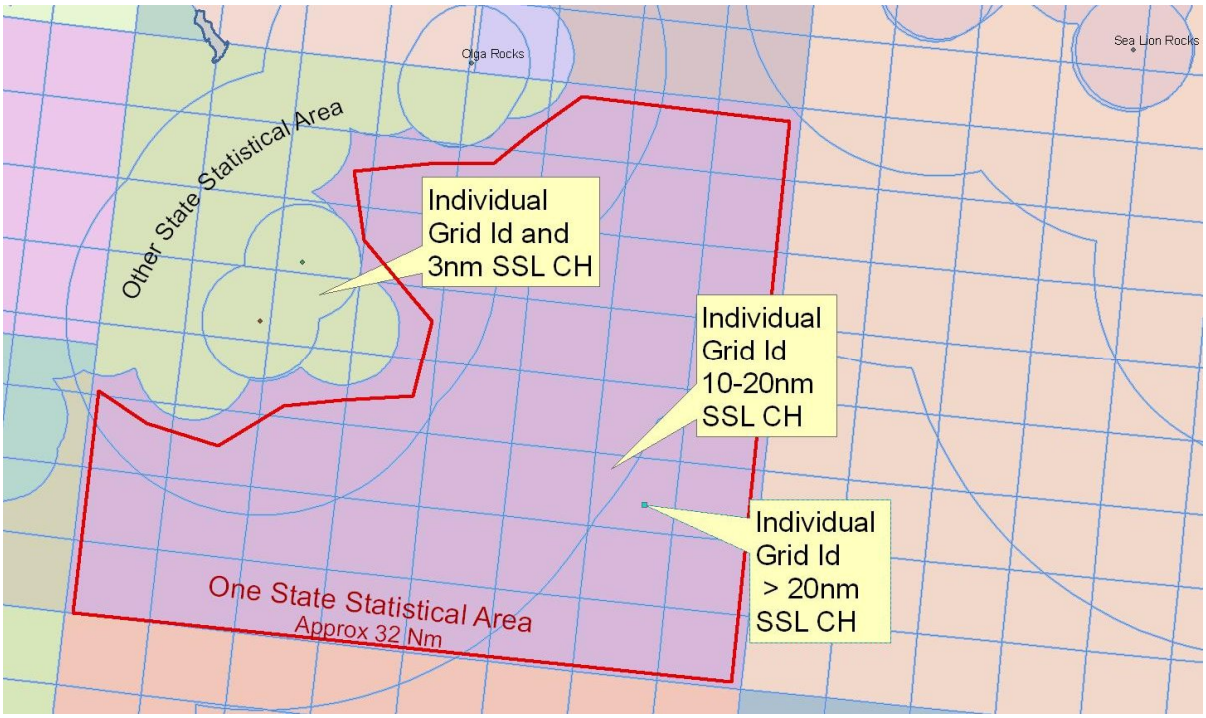
VMS-Observer Enabled Catch-In-Areas Database

Steve G. Lewis
GIS Coordinator/Analyst/DB
NOAA Fisheries, Alaska Region

In 2007, NMFS/Alaska Region began developing a fisheries harvest database that would integrate data acquired from onboard observers and data on vessel movements acquired by satellite through the Vessel Monitoring System (VMS). This VMS-Observer Enabled Catch-In-Areas (VOE-CIA) database is designed to increase the spatial resolution of the Catch Accounting System for both the observed and unobserved vessel fleet and thus to facilitate more accurate analysis of fisheries management issues.

The VOE-CIA database integrates catch data from the Catch Accounting System (which has the spatial resolution of a NMFS Reporting Area) into a database that resolves the GIS data into polygons with areas of approximately seven kilometers. In an unrestricted area, sixty four grid IDs fit inside one state statistical area. However, a given seven-kilometer polygon may be further divided into smaller polygons by the boundary of state statistical areas, the boundary of state and federal waters, or by the boundary of Steller sea lion critical habitat (broken out at 3, 10, and 20 nautical miles from one of 154 Steller sea lion rookeries or haulouts). Where confidentiality needs to be protected, a seven-kilometer polygon may be grouped with others into 20km polygons. Each polygon (the exact size of which will vary with latitude) and its subparts will have a distinct grid ID.



Splitting the Catch Accounting data from NMFS Reporting Areas into these grid IDs requires an iterative and ordered process; no single step can capture all the data. To start, a record is reported and entered into the database, and a unique transaction ID is created for that record. A record is considered either a single haul for an observed vessel, a single fishing trip for an unobserved catcher vessel, or a single week—as designated by the week-ending-date—for an unobserved catcher processor (at present, this is the finest temporal catch resolution currently available; in 2009, however, catcher processors will begin reporting at a finer temporal resolution).

After the transaction ID is established for that record, one of the following six steps is then used to incorporate the record into the Catch-in-Areas database. (Note that the following tables and figures use 2008 data solely for purposes of illustrating the operations of the database.)

1) The first step in the process coordinates the date and time of observed deployment and retrieval of gear with the vessel's VMS points that are within the same observed date and time. This 'fixes' the VMS points associated with an observed haul.

   VMS data are designed to transmit position reports every 30 minutes. It is probable that the process could miss the first and last VMS point by only a few minutes since it is based on Observed times. Therefore, a trackline is also drawn between the observed and deployed locations. A distinct set of grid IDs for both the VMS and Observer points are coordinated and associated.

   The associated grid IDs from the steps above are then attributed an equal amount of the catch for that record. Hence, a record that has eight grid IDs associated with it will receive 12.50% of the catch for that record from Catch Accounting.

   In 2008, 827,140 tons or 47.4% of the catch was matched in Step 1; and 52.6% of the catch remained to be matched in the processes that follow.

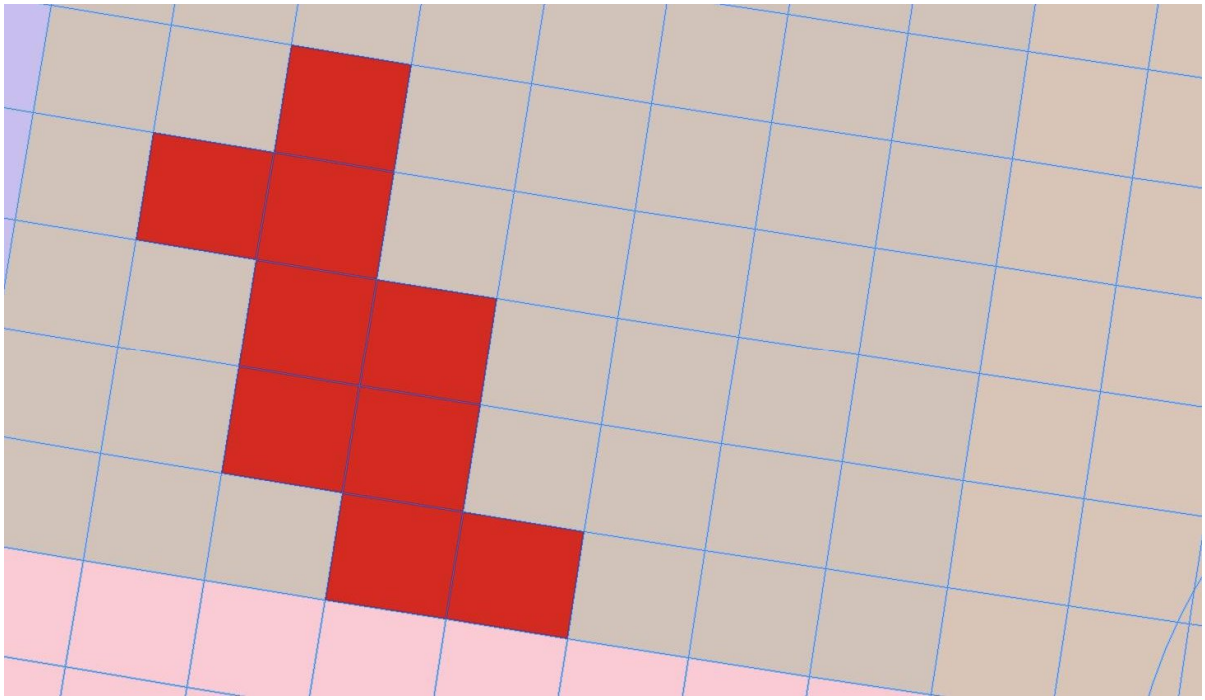
The tables below indicate average number of Grid IDs that were captured in Step 1: VMS-Observer by Date/Time matching process. The average is based on individual hauls shown by each row in the table. The data is shown in three base groups: FMP, FMP and harvest sector, and, FMP, harvest sector, and target fishery.

FMP	Avg#Grid IDs / Grid
AI	6
BS	8
GOA	6

FMP	Harvest Sector	Avg#Grid IDs per Grid
AI	CP	7
AI	CV	12
BS	CP	7
BS	CV	16
GOA	CP	6
GOA	CV	5

FMP	Harvest Sector	Example Species Code	Avg#Grid IDs per Grid
AI	CP	Pcod	7
AI	CP	Rock	3
AI	CV	Pcod	13
AI	CV	Rock	5
BS	CP	Pcod	9
BS	CP	Rock	4
BS	CP	Plck	5
BS	CV	Plck	17
GOA	CP	Pcod	7
GOA	CP	Rock	4
GOA	CV	Rock	4

A graphic illustrating captured Observed grid IDs (red - highlighted blocks below) from Bering Sea using a combination of VMS and Observer data.



2) The next step uses observer data that were not matched from Step 1. Some vessels are unmatched from Step-1 because transponder IDs may not be directly associated with a vessel ID for a given trip: for example, a vessel may lend a VMS transponder to another vessel, but the database fails to be updated to reflect that before catch is assigned to a trip/haul.

   As in the observer data process above, a line is drawn from the observer deployment location to the retrieved location, and the associated grid IDs are identified for that trackline. Catch is equally apportioned between the grid IDs for that record.

   In 2008, 219,709 tons or 12.59% of the catch was matched in Step 2; and 40.01% of the catch remained to be matched.

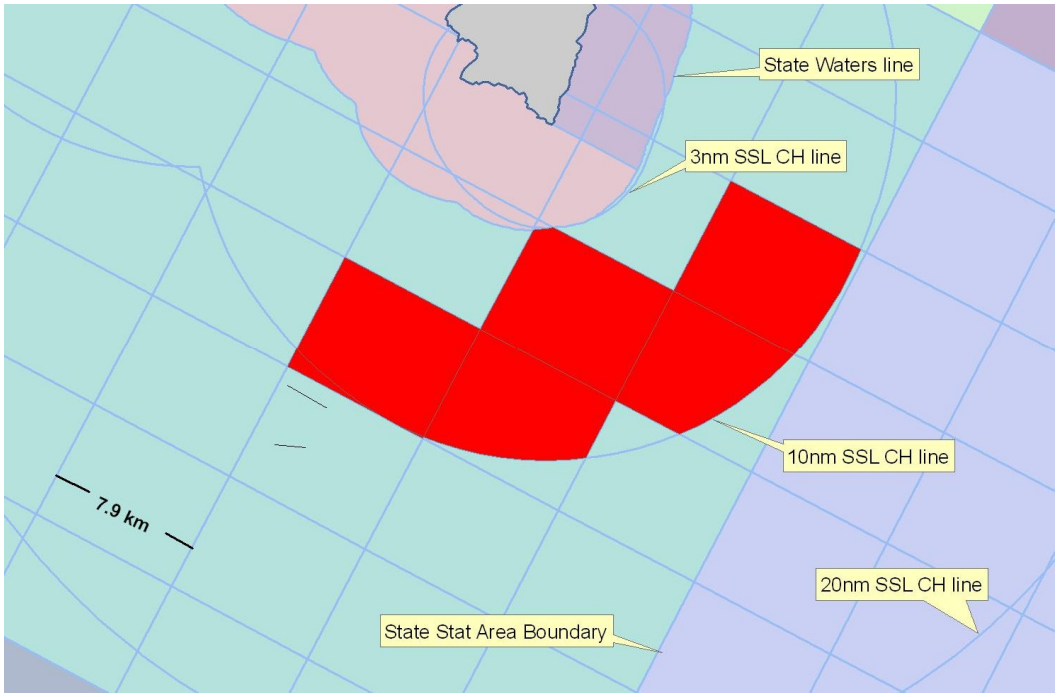
The tables below indicate average number of Grid IDs that were captured in Step 2: an individual observed haul trackline from observed deploy location to the retrieve location. The average is based on individual hauls shown by each row in the table. The data is shown in three base groups: FMP, FMP and harvest sector, and, FMP, harvest sector, and target fishery.

FMP	Avg#Grid IDs
AI	6
BS	8
GOA	5

FMP	Harvest Sector	Avg#Grid IDs
AI	CP	7
AI	CV	8
BS	CP	7
BS	CV	16
GOA	CP	5
GOA	CV	5

FMP	Harvest Sector	Example Species Code	Avg#Grid IDs
AI	CP	Pcod	8
AI	CP	Rock	3
AI	CV	Pcod	9
AI	CV	Rock	6
BS	CP	Pcod	9
BS	CP	Rock	4
BS	CP	Plck	7
BS	CV	Plck	16
GOA	CP	Pcod	7
GOA	CP	Rock	4
GOA	CV	Rock	5

A graphic illustrating captured Observed grid IDs (red - highlighted blocks below) that were not captured in Step 1.



3) The next step uses VMS data to capture an individual record for unobserved catcher vessels. In order to capture a vessel 'fishing,' four criteria must be in place: 1) A vessel must be operating between .9 knots and 4.1 knots; 2) a vessel must not be in an area known not to be a fishing area, e.g., very near ports; 3) a vessel must be operating inside at least one of the state statistical areas reported on its fish ticket; and 4) the date of the VMS point must match the date range on the fish ticket.

   We use the vessel's VMS points to calculate vessel speed for the database. In a GIS Albers conic coordinate system, we find the meters traveled using the Pythagorean Theorem and divide that by the time between one VMS point and the next.

   A catch record is weighted by how many VMS points are associated with a particular grid ID that met the four criteria above. For example, a vessel transiting through Unimak Pass: the vessel has to slow down to fishing speed (greater than .9 knots and less than 4.1 knots), is not in an area known not to be a fishing area, is inside at least one of the state statistical areas reported for the vessel, and has a trip time within the date range on the fish ticket. A single ping will be associated with that grid ID even though the vessel may not have been fishing. But a few hours later the vessel gets to its fishing grounds and continues to fish for the next two days. The vessel's trip time was three days. For two days (48 hours) the vessel met all of four of the criteria for fishing.

The single grid ID associated with Unimak Pass receives 1/48th (2.08%) of the catch. If the vessel spends a full day in one grid ID, that grid ID gets nearly 50% of the catch. If the vessel then spends the entire next fishing day equally in eight other grid IDs, each of those eight grid IDs gets 6.25% of the catch. It should be noted that this is a simple example and chances are that a vessel will not meet all four criteria for two full days.

A final adjustment is made after the catch is weighted. Consider a catcher vessel targeting flatfish in the GOA and which uses its MRA to top off with Pacific cod on the way back to port. On the fish ticket the vessel is reported to have been in one state statistical area with a catch composed of mostly flatfish and in another state statistical area with a catch of mostly Pacific cod. We do not reapportion the total amount of the catch; we only adjust the species composition in the grid ID associated with state statistical areas. This algorithm will not change the overall species composition or the overall catch weight associated with a grid ID.

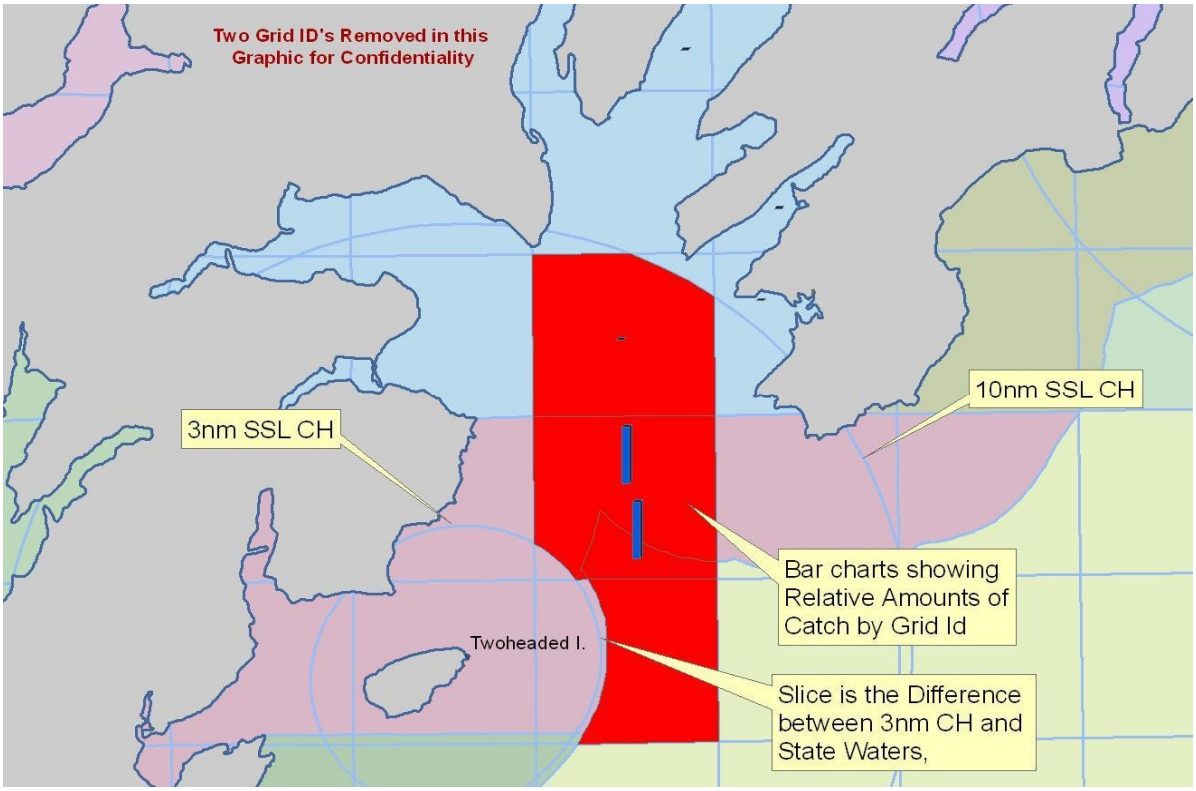
In 2008, 569,074 tons or 32.65% of the catch was matched in Step 3; and 7.35% of the catch remained to be matched in the following steps.

The tables below indicate average number of grid IDs that were captured in Step 3. The four criteria for the catcher vessel: speed, trip dates, fishing area, and state stat area. The average of captured grid IDs is based on individual trips. The data is shown in two base groups: FMP and FMP and target fishery.

FMP	Avg#Grid IDs
AI	15
BS	19
GOA	10

FMP	Harvest Sector	Example Species Code	Avg#Grid IDs
AI	CV	Pcod	9
AI	CV	Rock	14
AI	CV	Plck	7
BS	CV	Pcod	17
BS	CV	Plck	20
GOA	CV	Pcod	8
GOA	CV	Rock	9
GOA	CV	Plck	7

A graphic illustrating a catcher vessel's trip and the grid IDs captured using the criteria outlined in Step 3. Blue bar charts show relative amounts of catch distribution by grid ID. Captured grid IDs shown in red - highlighted blocks below



4) Some catcher vessels may not accurately report their state statistical areas. In step 4, we drop the requirement for state statistical areas and replace it with NMFS Reporting Areas. The four criteria become: 1) a vessel must be operating between .9 knots and 4.1 knots; 2) a vessel must not be in an area known not to be a fishing area, e.g., very near ports; 3) a vessel is operating inside their reported NMFS Reporting Areas; and 4) the date of the VMS point must match the date range on their fish ticket.

   As with Step 3, this catch is weighted as to how many VMS fishing points are associated with a Grid ID. No reapportionment of catch composition is completed in this step.

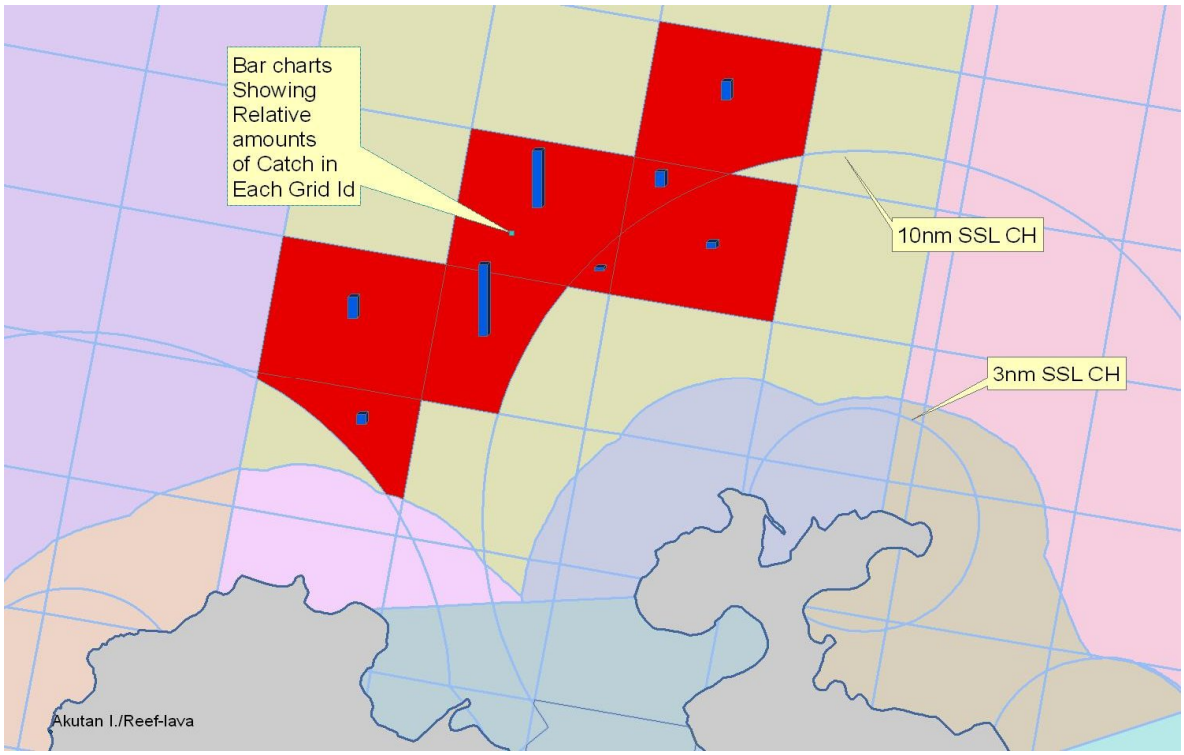
   In 2008, 20,683 tons or 1.19% of the catch was matched in Step 4.; and 6.17% of the catch remained to be matched in the following steps.

The tables below indicate average number of Grid IDs that were captured in Step 4. The 4 criteria for the catcher vessel: speed, dates, fishing area, NMFS Reporting Areas. The average is based on individual trips. The data is shown in two base groups: FMP and FMP and target fishery.

FMP	Avg#Grid IDs
AI	11
BS	13
GOA	8

FMP	Harvest Sector	Example Species Code	Avg#Grid IDs
AI	CV	Pcod	6
BS	CV	Pcod	10
BS	CV	Plck	16
GOA	CV	Pcod	8
GOA	CV	Rock	7
GOA	CV	Plck	8

A graphic illustrating a catcher vessel's trip. Grid IDs captured using the criteria outlined in Step 4. Blue bar charts showing relative amounts of catch based on time the vessel spent inside Grid IDs. Captured grid IDs shown in red - highlighted blocks below.



5) Step 5 addresses unobserved catcher processors who report weekly on their production. Like an unobserved catcher vessel without a state statistical area, four criteria must be met: 1) A vessel must be operating between .9 knots and 4.1 knots; 2) a vessel must not be in an area known not to be a fishing area, e.g., very near ports; 3) a vessel must be operating inside its reported NMFS Reporting Areas; and 4) the date of the VMS point must match the week ending date reported on the catcher processor's weekly production report. In 2009 with additional reporting for unobserved catch processors, the temporal resolution will increase and hence the data for this step. Additionally, some catcher vessels are captured in this step by week ending date rather than by their reported trip dates.

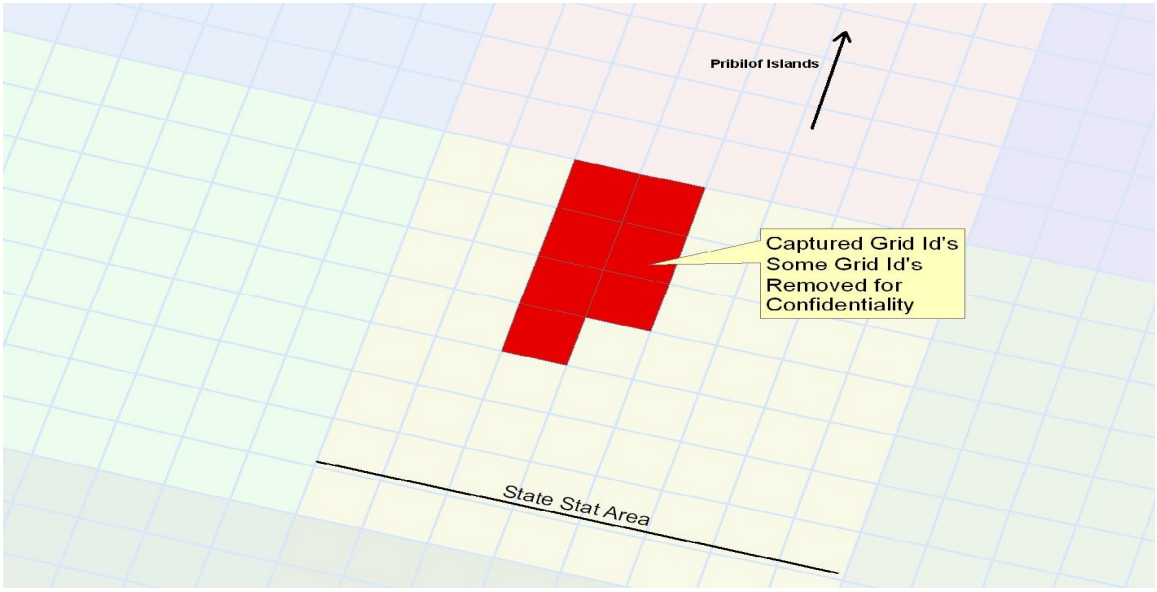
The tables below indicate average number of grid IDs that were captured in Step 5: The four criteria for these unmatched unobserved vessels: speed, week ending date (Saturday), fishing area, and NMFS Reporting Area. The average is based on a week ending date. The data is shown in three base groups: FMP, FMP and harvest sector, and, FMP, harvest sector, and target fishery.

FMP	Avg#Grid IDs
AI	3
BS	4
GOA	3

FMP	Harvest Sector	Avg#Grid IDs
AI	CP	3
AI	CV	2
BS	CP	4
BS	CV	2
GOA	CP	3
GOA	CV	2

FMP	Harvest Sector	Example Species Code	Avg#Grid IDs
AI	CP	Pcod	3
AI	CV	Pcod	2
AI	CV	Plck	2
BS	CP	Pcod	4
BS	CP	Plck	3
BS	CV	Pcod	2
BS	CV	Plck	2
GOA	CP	Pcod	3
GOA	CP	Rock	2
GOA	CV	Pcod	2
GOA	CV	Rock	2
GOA	CV	Plck	2

A graphic illustrating an unobserved weekly trip. These grid IDs were captured using the criteria outlined in Step 5. Captured grid IDs shown in red - highlighted blocks below. Some grid IDs were removed for confidentiality.



Steps 1 through 5 above capture 96.13% (for the 2008 data) of the catch from Catch Accounting inside one of the seven-kilometer grid IDs. The final steps, called Average Vessel, match catch from the previously matched vessels (from steps 1 – 5) to the unmatched vessel records. All but 604 tons (for the 2008 data) of the unmatched catch are matched using this final process.

6) The Average Vessel algorithm groups all previously matched vessels operating in the groupings shown below, and then apportions catch equally to the associated grid IDs for the unmatched records. The first grouping includes vessel ID. Vessel ID is included with week ending date, NMFS Reporting Area, Harvest Sector, Gear, Target, etc., as we assume the best extrapolation is on a vessel operating as itself. We have seen this grouping to be effective when a catcher vessel with multiple trips in a single week may not be captured during a single trip due to a reporting or recording error.

   The following groupings, shown in the table below, were coordinated by such aspects as Management Program Code, Harvest Sector, NMFS Reporting Area, Gear, Target, and Week Ending Date. After matches for all those groupings are found (between the unmatched records in catch accounting and the previously match records in Catch-In-Areas), the grid IDs are compiled for those matched records and the catch is evenly divided among those grid IDs.

After an average vessel record is apportioned to a set of grid IDs, a transaction ID is created and that vessel record is removed from further matching. The groupings for Average Vessel are then slightly liberalized, and the next groupings are formed, matched and apportioned to grid IDs. As noted above, these steps capture greater than 99.98% of the catch. Catch that is not captured is often groundfish caught by non-federally permitted groundfish catcher vessels.

Match-Groupings for the Iterative Average Vessel Extrapolation Algorithm.

- Mgt_Prog_Code HarvestSector Rpting Area Target, Gear WeekEndDate Vessel ID
- Harvest Sector NMFS Area Gear Target WeekEndDate Processor ID
- Mgt_Prog_Code HarvestSector NMFS Area Gear Target WeekEndDate
- Mgt_Prog_Code HarvestSector NMFS Area Gear WeekEndDate Target
- Mgt_Prog_Code NMFS Area Gear Target WeekEndDate
- Harvest Sector NMFS Area Target WeekEndDate
- Harvest Sector NMFS Area Gear WeekEndDate
- NMFS Area Gear Target WeekEndDate
- NMFS Area Target WeekEndDate
- NMFS Area Gear WeekEndDate
- NMFS Area Gear Target Month Year
- NMFS Area Target Month Year
- NMFS Area Gear Month Year
- FMPAreaCode Gear Target WeekEndDate
- FMPAreaCode Target WeekEndDate
- FMPAreaCode Gear WeekEndDate
- FMPAreaCode Gear Target Month Year
- FMPAreaCode Target Month Year
- FMPAreaCode Gear Month Year

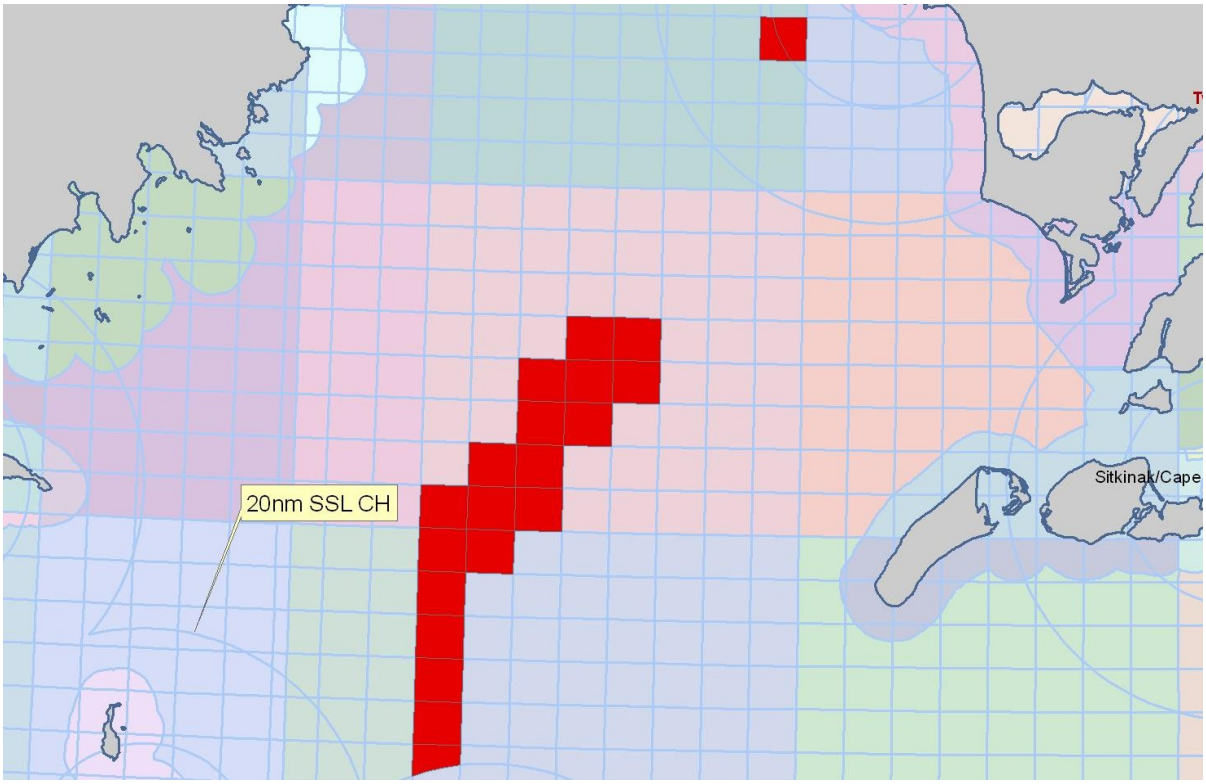
For clarity, the following summary tables aggregate all 19 levels of the Average Vessel extrapolation algorithm into a single set of tables.

FMP	Avg#Grid IDs
AI	33
BS	32
GOA	28

FMP	Harvest Sector	Avg#Grid IDs
AI	CP	36
AI	CV	23
BS	CP	38
BS	CV	30
GOA	CP	33
GOA	CV	28

FMP	Harvest Sector	Example Species Code	Avg#Grid IDs
AI	CP	Pcod	36
AI	CV	Pcod	23
BS	CP	Pcod	39
BS	CP	Plck	24
BS	CV	Pcod	33
BS	CV	Plck	30
GOA	CP	Pcod	34
GOA	CP	Rock	27
GOA	CV	Pcod	28
GOA	CV	Rock	28
GOA	CV	Pcod	13
GOA	CV	Plck	2

This graphic illustrating the Average Vessel Extrapolation Algorithm grid id's that was captured, shown in red - highlighted blocks below. This Average Vessel was grouped and matched on a vessel or group of vessels with the same Harvest Sector, NMFS Reporting Area, Gear Type, Target, and Week Ending Date.



The table below illustrates the amount of catch by each matching method.

Analysis based on 2008			
Matching Method	**Tons Matched**	**% of Total Catch**	**Cumulative % Matched**
VMS-Obs by Time and Obs Trackline	827,140	47.39%	47.39%
OBS Deploy and Retrieve Trackline	219,709	12.59%	59.98%
CV-Stat_Area	569,754	32.65%	92.63%
CV-NMFS_Area	20,683	1.19%	93.82%
CP_NMFS_Area	40,332	2.31%	96.13%
Grouping for Extrapolations for unmatched catch:			
Avg_MgtPrg_HS_RA_Gr_Tgt_WED_Ves	1,321	0.08%	96.20%
Avg_HS_RA_Gr_Tgt_WED_VesID	24	0.00%	96.20%
Avg_HS_RA_Gr_Tgt_WED_PID	32,466	1.86%	98.07%
Avg_MgtPrg_HS_RA_Gr_Tgt_WED	17,701	1.01%	99.08%
Avg_MgtPrg_RA_Gr_Tgt_WED	513	0.03%	99.11%
Avg_HS_RA_Tgt_WED	5,829	0.33%	99.44%
Avg_HS_RA_Gr_WED	4,516	0.26%	99.70%
Avg_RA_Gr_Tgt_WED	166	0.01%	99.71%
Avg_RA_Gr_WED	447	0.03%	99.74%
Avg_RA_Tgt_WED	250	0.01%	99.75%
Avg_RA_Gr_Mnt_Yr	2,534	0.15%	99.90%
Avg_FMP_GrT_Tgt_WED	894	0.05%	99.95%
Avg_FMP_Gr_Mnt_Yr	16	0.00%	99.95%
Avg_FMP_Tgt_WED	582	0.03%	99.98%
Avg_FMP_Gr_WED	23	0.00%	99.98%
Total VOE-CIA by Grid_ID to Catch Accounting	1,744,900		
Total of full Catch Accounting System	1,745,504		

The final dataset includes data from Steps 1 – 5 above, plus data derived from the Average Vessel processes. This creates a geospatial database that matches the Catch Accounting system. Several additional columns of information are added to Catch Accounting that include Percent in Grid, Weight-In-Grid, Match Source, 'ESA Critical Habitat,' '679 Critical Habitat,' and assorted protection areas. Each area of study resides in a separate column (which may be queried) to insure that catch is not double or triple counted.

Match Source is the metadata column. It provides analysts information as to which step captured the data: Step 1: VMS-Obs, Step 2: OBS, Step 3: CV-Stat_Area, Step 4: CV-NMFS_Area, Step 5: CP_NMFS_Area, or Average Vessel. Average Vessel is further broken down by which groupings were used for the extrapolations. For instance, the first grouping above includes AVG: Harvest Sector-NMFS_Area GEAR Type, Target, Week Ending Date and Vessel Id. The Average Vessel catch can be removed from queries if requested by the analyst.

With the database complete, it can then be joined back to the GIS, or a GIS feature class can be joined to the native database by the grid ID. Other geospatial data that are currently complete and attached to the CIA include distance from aggregated Steller sea lion Critical habitat sites; distance from individual, overlapping SSL sites; and distance from foraging areas and some of the habitat protection and conservation areas.

This table illustrates most of the relevant columns in the VOE-CIA dataset. Note that data can be selected independently or grouped by any of the columns bellow, including, Target Fishery, Gear Type, Vessel ID, Processor, Sector, Management Program, Coop or Group or operating in any of several zones (SSL or Habitat) or management areas.

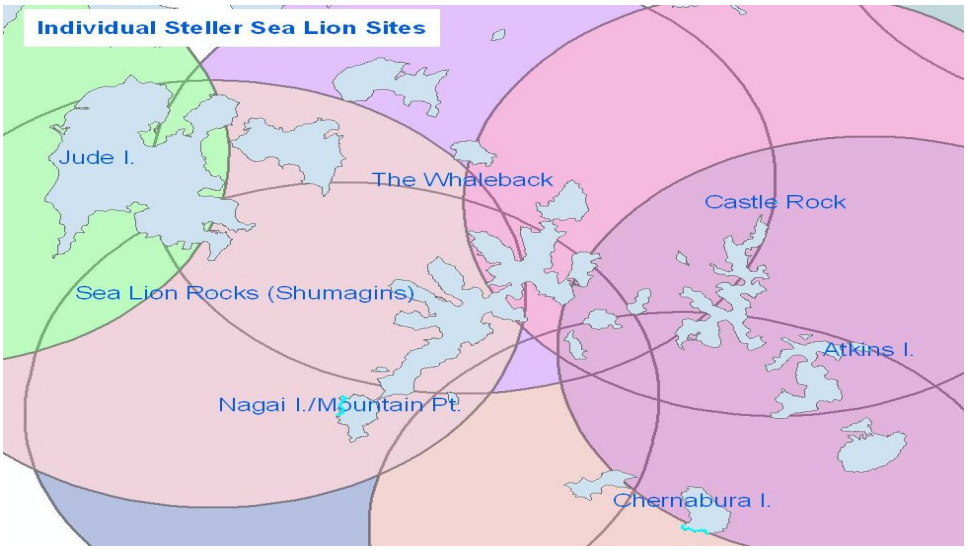
Base Catch Accounting Data
Reporting Area Code
Catch Activity Date
Week End Date
Trip Target Date
Year, Month, Quarter
Catch Report Type Code
CA Reference / Haul-SLog Join
Vessel ID
Gear Type
Harvest Sector
Trip Target Code
Management Program Code
AFA Coop ID
Processor ID
State Waters Flag
FMP Area Code
Species Group Code
BSAI Processing Sector
Vessel Size Catagory
PSCNQ Processing Sector
CDQ Group ID
Agency Species Code
Source Table: Obs, WPR, State
Directed Fishing Flags
Weight Posted

Additional VOE-CIA Columns
7Km Grid ID
Weight In Grid
Match Source: Matching Algorthm
Species Adjusted Weight
ADFG STAT AREA
Percent in Grid
20Km Grid ID
226 SSL Critical Habitat
679 SSL Critical Habitat
No NPT Areas

Other Distinct VOE-CIA Datasets
Overlapping SSL Sites
PSC: Prohibited Species

Other Datasets: Prohibited Species and Overlapping Steller Sea Lion Site VOE-CIA Datasets

Two separate VOE-CIA datasets have also been created: Prohibited Species (PSC) and Overlapping SSL sites. The overlapping SSL site dataset is by each of the 154 Steller sea lion sites, split out by 3, 10 and 20nautical miles; and, where the individual SSL sites overlap, the catch will overlap. This will give analysts and policy makers the ability to look at individual vessels, fleets, and target fisheries, gears types etc., operating in or around each individual SSL sites. Catch by the overlapping Steller sea lion site cannot be grouped and summed by management areas since catch from the overlapping Steller sea lion sites would be counted several times where the sites overlap.



PSC: The PSC database (PSC) is joined by the associated values to the VOE-CIA and the records divided into Grid ID's in the same proportions that were made with Catch Accounting groundfish database. The noted caveats to this PSC dataset are embedded within the PSC data. These caveats include how the base PSC data was collected and then extrapolated to the non observed fleets.

Included Prohibited Catch Species.

Blue King Crab
Bairdi Tanner Crab
Chinook
Grenadier
Hake
Golden King Crab
Herring
Halibut
Non Chinook Salmon
Other King Crab
Red King Crab

Use of the VOE-CIA for Analytical Purposes

The VOE-CIA database uses an iterative, ordered process to match VMS records, Observer collected data and VMS/Catch Accounting System indicators to a fishing vessel. This gives analysts the capability to analyze unobserved vessels that may have been transparent when only using earlier analytical tools such as observer data. For example, comparative analysis shows a difference in catch between the VOE-CIA and the Expanded Observer Dataset (extrapolated Observer data, also called the EOD) for the unobserved/small vessel fleet that operates within 3 and 10 nm from unrestricted Steller sea lion sites.

It should be noted that VOE-CIA data only go back as far as 2003. This is due to the unavailability of reliable VMS data and a vessel linked catch accounting system for 2003. Observer data on the other hand goes back to the early 1990s, giving analysts the ability to look at long-term trends in groundfish catch and can relate it to Steller sea lion population trends. Both VOE-CIA and the EOD are utilized in this document to insure the best available data is being used for the appropriate analysis.