

SEFSC Miami Laboratory's Pelagic Observer Program vessel sampling/selection as described in Christman (2014).

The entire study area is stratified into 11 geographic regions and 4 calendar quarters. For each stratum (hereafter a georegion \times quarter combination), a list of fishing vessels and associated effort is generated based on the logbook reports of number of sets (effort) the previous year. This constitutes the sampling frame from which vessels are selected to receive observers this year. Vessels are associated with only one georegion per quarter but can appear in every quarterly spatial sampling frame. Vessels are selected with probability proportional to their previous year's effort in the georegion and it is possible for a vessel to be selected in more than one quarter.

Sample size within each stratum is based on the previous year's total fishing effort (number of sets) in that stratum. Hence, the number of vessels selected for observer coverage must have had cumulative fishing effort that sums to at least 8% of the previous year's total effort in that stratum. As a result, the actual number of vessels selected is a function of the random selections using probability proportional to effort.

Once selected, observers are placed onboard selected trips taken by the vessel until a minimum number of sets are observed in order to ensure that the stratum-level sample coverage is at least 8% of the previous year's total effort. Once a vessel is selected, the next decision is which fishing trips may include an observer on board. It is intended that a minimum number of sets (C) be observed for each selected vessel, so preference is given to trips of sufficient length that the minimum number of sets can be observed in a single trip. This maximizes the time an observer is actively collecting data from sets. This may not be problematic for selection of trips since many trips tend to be of similar length within each area and quarter. Because of the vagaries of weather or mechanical issues precluding some intended trips from occurring, trips also may be selected for observer coverage according to time order, with earlier trips (within the quarter) more likely to be selected than later trips.

The final decision, which sets to observe once onboard a vessel, is that all sets are observed regardless of the total number of sets deployed on the trip. Occasionally, a set may be interrupted and is then considered to be two sets for purposes of recording effort.

This sampling strategy can be described as a non-probabilistic stratified 3-stage cluster sampling design in which at the first stage a cluster (vessel) is selected proportional to size and without replacement and at the second stage subunits (trips) are sampled non-randomly, or approximately according to size (expected # sets on the trip). The third stage, selections of sets, is non-random and observed with probability 1.

Christman, Mary C. 2014. Report on Model-based versus Design-based Approaches to inference in Observer Programs. MCC Statistical Consulting LLC. Final Report to the Fisheries Statistics Division, NOAA/NMFS/SEFSC Miami FL.