INTERNATIONAL DOLPHIN CONSERVATION PROGRAM

SCIENTIFIC ADVISORY BOARD

5TH MEETING

CANCUN (MÉXICO)
21 JUNE 2007

DOCUMENT SAB-05-03a

ESTIMATION OF MISREPORTED MORTALITY

The IATTC staff has been carrying out investigations aimed at detecting observer misreporting of dolphin mortality. This has included documentation of perceptions about observers by field office staff, examination of data by editors, and the use of statistical techniques to highlight unusual reports that might indicate fabrication of data. This paper outlines procedures that could be used to estimate misreported mortality and briefly discusses several points regarding existing and future data requirements.

Statistical procedure for estimating misreported mortality

The following is an outline of the procedure that could be used to develop a preliminary estimate of unreported dolphin mortality in dolphin sets; a related approach could also be considered for estimating missing mortality in dolphin sets that were reported as unassociated sets. This procedure would consist of five steps: i) identification of observers whose data are suspect; ii) identification of specific dolphin sets of these observers with suspect mortalities; iii) evaluation of the efficiency of steps (i)-(ii) to identify suspect data under known modifications to mortalities; iv) estimation of mortality for sets identified in step (ii); and v) evaluation of the sensitivity of the estimate of step (iv) to parameters used in steps (i)-(ii). Steps (i)-(ii) represent the approach currently used to review observer data[1,2].

i. Identification of observers whose data are to be considered suspect

An estimate of the probability that an observer would be associated with as many or more suspect sets by chance will be used to define which observers will be considered to have provided suspect data (see references above for further details). It is recommended that the threshold for this probability be set at 0.01 to start. That is, only the data of observers who had less than a 1% chance of having as many or more suspect sets will be considered further. This threshold is one of two parameters used in the data screening algorithm that may affect the estimate of misreported mortality (see (v) below). Additional data that could be collected to help bound the choice of these parameters is discussed below (see (vi)a below).

ii. Identification of sets with misreported dolphin mortality data

All dolphin sets of observers identified in (i) that had no reported dolphin mortality, but for which dolphin mortality would have been predicted, will be considered to have misreported mortality data. In terms of the data screening algorithm, these would be any sets reported by observers identified in (i) with a residual ≤ -0.5.

iii. Simulations to evaluate efficiency of data screening algorithm

Simulations will be conducted to evaluate efficiency of the data screening algorithm under various known types of modifications to the mortality data. An appropriate subset of the observer data will be identified for the simulations, most likely data of those observers for whom no suspect data were identified in (i). The reported mortality in sets of these observers will be modified (*e.g.*, sets with mortalities greater than 10 animals will be replaced with 0's) and then screened as per (i)-(ii) above. The percentage of suspect sets and observers with suspect data correctly identified will be determined. The data screening algorithm

[1] Lennert-Cody, C.E. and Berk, R.A. 2007. Statistical learning procedures for monitoring regulatory compliance: an application to fisheries data. Journal of the Royal Statistical Society Series A 170 Part 3: 1-19.
[2] Document IRP-39-08b (http://www.iattc.org/PDFFiles2/IRP-39-08b-Measurement-of-performance.pdf)

will be evaluated on several different types of modifications to the data (*e.g.*, replacing high mortality sets with 0's; reducing mortality by a fixed percent).

iv. Methods for estimating dolphin mortality of sets with suspect data

Several different methods will be compared for estimating dolphin mortality in sets identified in (ii): average mortality per set, average mortality per ton of yellowfin tuna caught, stochastic gradient boosting, and random forest/quantile random forests. The performance of these five methods will be compared by first constructing an estimator on an appropriate training data set, and then estimating the mortality on an appropriate test data set with a known mortality. The method selected will be the one whose estimate of mortality is closest to the known mortality of the test data set.

An alternative would be to consider all data of observers identified in (i) as suspect, and estimate missing mortality for all sets of these observers, instead of restricting estimation to sets where no mortality was reported but mortality was predicted. This is similar to the procedure carried out for the calculation of vessel performance (see above references). The results of these two approaches will be compared.

v. Sensitivity of estimated mortality to data screening parameters

The procedure described in (i)–(iv) will provide an estimate of the amount of misreported dolphin mortality in dolphin sets. To obtain credible bounds for such an estimate, independent data will be needed to establish credible bounds on the parameters used in the data screening algorithm, and it is recommended that the collection of such data be discussed at a meeting of Observer Programs (see (vi) below). For the present, the parameters used in the data screening algorithm will be varied, and estimates of misreported mortality computed in order to determine the sensitivity of the estimated mortality to these parameters.

vi. Data considerations

To estimate the amount of misreported dolphin mortality for the international fleet, it is essential that the data of all AIDCP observer programs be screened. It is recommended that, as part of the data-screening process, programs exchange information on vessels and captains associated with suspect data.

a. Additional data to improve data screening algorithm

Additional data should be collected that can be used to improve the data screening algorithm. For example, information on diver activities both during and after backdown (*e.g.*, time in the water, location within the net), and characteristics of brailing activities would likely prove useful. The types of additional data that could be collected should be discussed at a meeting of Observer Programs.

b. Bounding parameters used in the data screening algorithm

In order to bound the choice of parameters used in the data screening algorithm, independent data are needed. For example, one such source that should be considered is data on vessel trips not taken by observers. An analysis of data on the temporal patterns in observers available to take trips, and vessels awaiting observers, separate from trips actually taken, may provide information on vessel preferences for particular observers. Existing data sources that are not presently being used for data screening include the IRP data base on infractions, which might be used to develop an index of the likelihood of a captain or vessel being involved in attempts to have observers misreport data. The types of data which might be collected (or currently exist, but are not used) to provide independent information on observers who may be misreporting data should be discussed at a special workshop of the Scientific Advisory Board.