

INTERNATIONAL DOLPHIN CONSERVATION PROGRAM

INTERNATIONAL REVIEW PANEL

40TH MEETING

LA JOLLA, CALIFORNIA (USA)

19 OCTOBER 2005

DOCUMENT IRP-40-12

COMPARISON OF OBSERVER PROGRAMS

This paper presents comparisons between the IATTC observer program, by country, and the Programa Nacional de Observadores Pesqueros de Ecuador (PROBECUADOR) for Ecuadorian-flag vessels, the Programa Nacional de Aprovechamiento del Atún y Protección de Delfines (PNAAPD) for Mexican-flag vessels, and the Programa Nacional de Observadores de Venezuela (PNOV) for Venezuelan-flag vessels. Comparisons between the IATTC program and the Programa Nacional de Observadores de Túnidos (PNOT) for Spanish-flag vessels and for United States vessels covered by the Forum Fisheries Agency (FFA) were not performed due to insufficient data. The average differences between programs for 2000-2004 shown in items 2, 3, 5, 6 and 7 were tested statistically using a randomization test, as outlined in Appendix A, when data were available for three or more years for both programs. Statistical comparisons of other items were not possible because not all national programs provide by-set data to the IATTC. Average differences between programs were considered *significant* if the statistical test yielded a probability of 0.05 or less.

'Calendar year data' means data for sets that took place between January 1 and December 31 of a particular year. For example, for a trip that started in November 1999 and ended in February 2000, only data from sets made on or after January 1, 2000, would be used in the analyses for calendar year 2000. 'Departure year data' means that all data from trips that departed in a particular year were used; thus, all the data of a trip that departed in November 2000 and finished in February 2001 would be included in the analyses for 2000.

To streamline the comparison between programs, three previously presented comparisons (spatial distribution of sets by set type; percentage of tuna catch, by species; and average catch per day) have been eliminated because they were either ineffective at summarizing differences between programs and/or because the utility of the information provided was questionable.

1. PERCENTAGE OF SETS BY SET TYPE

Calendar-year data. National program data are from annual summaries provided to the IATTC by the national programs. IATTC data are from the IATTC permanent database. Accidental dolphin sets are treated as either unassociated sets or floating-object sets. Because of rounding, percentages may not sum to exactly 100%.

	Dolphin	Unassociated	Floating object	Whale	Dolphin	Unassociated	Floating object	Whale
	Program 1				IATTC			
2000	80.8	16.1	2.9	0.2	84.7	13.6	1.7	0.04
2001	86.5	10.7	2.8	0.05	88.0	10.0	1.8	0.1
2002	89.1	9.2	1.7	0.0	90.2	8.1	1.7	0.04
2003	86.7	10.1	3.1	0.1	85.9	10.5	3.3	0.3
2004	76.8	19.9	3.2	0.0	79.4	16.2	4.4	0.05
	Program 2				IATTC			
2000	58.0	38.3	3.3	0.4	65.1	30.5	4.1	0.2
2001	82.0	16.8	1.2	0.0	75.2	23.3	1.4	0.0
2002	87.7	11.3	1.0	0.0	87.4	12.1	0.5	0.0

2003	84.0	15.0	0.9	0.0	88.0	10.8	1.2	0.0
2004	75.2	22.8	1.9	<0.01	74.0	23.3	2.6	0.08
	Program 3				IATTC			
2000	2.0	12.0	86.0	0.0	0.0	16.8	83.2	0.0
2001	0.0	23.0	77.0	0.1	0.03	20.1	79.8	0.1
2002	4.9	26.0	69.0	0.1	3.7	29.7	66.3	0.2
2003	6.8	44.7	48.5	0.0	2.6	45.0	52.3	0.1
2004	2.9	41.6	55.4	0.06	5.4	43.3	51.1	0.13

2. PERCENTAGE OF TRIPS INVOLVING NO SETS ON DOLPHINS

Departure-year data, from the IRP database. The difference in the percentage of trips between programs was computed as the value for the national program less that for the IATTC. Excludes accidental sets. NS = not significant; S = significant.

	Program 1	IATTC	Difference	Average difference
2000	0.0	0.0	0.0	
2001	2.8	7.0	-4.2	
2002	2.6	0.0	2.6	
2003	4.4	3.2	1.2	
2004	5.1	5.7	-0.6	-0.2 (NS)
	Program 2	IATTC		
2000	6.7	4.3	2.4	
2001	1.2	6.3	-5.1	
2002	5.7	6.7	-1.0	
2003	10.5	8.5	2.0	
2004	9.9	10.8	-0.9	-0.5 (NS)
	Program 3	IATTC		
2000	83.3	100.0	-16.7	
2001	98.2	98.7	-0.5	
2002	97.1	94.8	2.3	
2003	94.0	96.0	-2.0	
2004	94.9	92.3	2.6	-2.9 (NS)

3. AVERAGE NUMBER OF DAYS PER TRIP

Departure-year data. Length of the trip is computed as the number of days from departure date to arrival date. Data are from the IRP database. The difference in the average number of days per trip between programs was computed as the value for the national program less that for the IATTC. NS = not significant; S = significant.

	Program 1	IATTC	Difference	Average difference
2000	53.7	46.7	7.0	
2001	34.2	33.0	1.2	
2002	32.8	33.6	-0.8	
2003	44.4	45.4	-1.0	
2004	54.7	50.7	4.0	2.1 (NS)
	Program 2	IATTC		
2000	49.0	48.3	0.7	
2001	41.0	39.8	1.2	
2002	34.2	34.4	-0.2	

2003	37.5	34.7	2.8	
2004	45.0	40.4	4.6	1.8 (NS)
	Program 3	IATTC		
2000	43.0	44.8	-1.8	
2001	49.9	47.3	2.6	
2002	44.8	44.8	0.0	
2003	38.5	38.4	0.1	
2004	41.2	41.3	-0.1	0.1 (NS)

4. PERCENTAGE OF INTENTIONAL DOLPHIN SETS WITH ZERO MORTALITY

Calendar-year data. Excludes accidental set data. National program summary data were provided to the IATTC by the national programs. The difference in the percentage of intentional dolphin sets with zero mortality between programs was computed as the value for the national program less that for the IATTC. IATTC data are from the IATTC permanent database. Dashed lines indicate no dolphin sets were reported.

	Program 1	IATTC	Difference
2000	92.5	91.7	0.8
2001	93.5	91.7	1.8
2002	93.9	93.5	0.4
2003	93.1	94.1	-1.0
2004	92.8	92.9	-0.1
	Program 2	IATTC	
2000	91.6	89.6	2.0
2001	92.5	91.0	1.5
2002	93.7	92.6	1.1
2003	94.0	93.8	0.2
2004	93.7	93.5	0.2
	Program 3	IATTC	
2000	100.0	-----	-----
2001	-----	100.0	-----
2002	97.3	99.1	-1.8
2003	99.1	94.4	4.7
2004	95.7	95.8	-0.1

5. AVERAGE MORTALITY PER SET

Calendar-year data. The average mortality per set (MPS) is computed as the sum of all dolphin mortalities in intentional dolphin sets, divided by the sum of all intentional dolphin sets during the study periods. The difference in the average mortality per set between programs was computed as the value for the national program less that for the IATTC. Data are from the IRP database. NS = not significant; S = significant. Dashed lines indicate no dolphin sets were reported.

	Program 1	IATTC	Difference	Average difference
2000	0.149	0.171	-0.022	
2001	0.228	0.172	0.056	
2002	0.132	0.117	0.015	
2003	0.130	0.117	0.013	
2004	0.140	0.161	-0.021	0.008 (NS)
	Program 2	IATTC		
2000	0.162	0.197	-0.035	

2001	0.143	0.151	-0.008	
2002	0.132	0.140	-0.008	
2003	0.105	0.106	-0.002	
2004	0.108	0.106	0.002	-0.010 (NS)
	Program 3	IATTC		
2000	0.0	-----	-----	
2001	-----	0.0	-----	
2002	0.013	0.018	-0.005	
2003	0.018	0.244	-0.226	
2004	0.149	0.125	0.024	-0.069 (NS)

6. AVERAGE RATE OF POSSIBLE OBSERVER INTERFERENCE INFRACTIONS

Departure-year data. The average rate of possible observer interference infractions is computed as the sum of the number of cases of interference reported by observers divided by the number of trips. The difference in the average rate of possible observer interference infractions between programs was computed as the value for the national program less that for the IATTC. Data are from the IRP database. NS = not significant; S = significant.

	Program 1	IATTC	Difference	Average difference
2000	0.030	0.0	0.03	
2001	0.0	0.0	0.0	
2002	0.0	0.0	0.0	
2003	0.0	0.016	-0.016	
2004	0.0	0.0	0.0	0.003 (NS)
	Program 2	IATTC		
2000	0.0	0.065	-0.065	
2001	0.0	0.025	-0.025	
2002	0.01	0.03	-0.02	
2003	0.0	0.0	0.0	
2004	0.0	0.0	0.0	-0.023 (S)
	Program 3	IATTC		
2000	0.0	0.053	-0.053	
2001	0.0	0.0	0.0	
2002	0.0	0.0	0.0	
2003	0.0	0.012	-0.012	
2004	0.0	0.013	-0.013	-0.015 (NS)

7. AVERAGE RATE OF POSSIBLE PROCEDURAL INFRACTIONS

Departure-year data. The average rate of possible procedural infractions (explosives use, night sets, no backdown, sets after reaching the DML, sacking up or brailing live dolphins, unavoided dolphin injury or death, and sets without continued rescue) is computed as the sum of such infractions reported by observers divided by the sum of the number of intentional dolphin sets. The difference in the average rate of possible procedural infractions between programs was computed as the value for the national program less that for the IATTC. Data are from the IRP database. NS = not significant; S = significant. Dashed lines indicate no dolphin sets were reported.

	Program 1	IATTC	Difference	Average difference
2000	0.062	0.089	-0.027	
2001	0.040	0.053	-0.013	

2002	0.018	0.019	-0.001	
2003	0.012	0.014	-0.001	
2004	0.009	0.011	-0.002	-0.009 (NS)
	Program 2	IATTC		
2000	0.006	0.045	-0.039	
2001	0.005	0.013	-0.008	
2002	0.003	0.007	-0.004	
2003	0.001	0.002	-0.001	
2004	0.004	0.002	0.002	-0.010 (S)
	Program 3	IATTC		
2000	1.0	-----	-----	
2001	-----	0.062	-----	
2002	0.086	0.0	0.086	
2003	0.0	0.031	-0.031	
2004	0.017	0.006	0.011	0.022 (NS)

The comparisons presented in this series of papers have focused on differences among the results of the programs. Any such differences do not necessarily indicate whether the results of one program are better than those of another, but may indicate persistent differences in observer training, data handling techniques, *etc.* Obviously, the programs should collaborate to try to identify and to remove any such differences. However, whether or not there are differences among programs, individual observers in any program could be reporting unreliable data, and a lack of differences between programs does not guarantee data reliability.

The Secretariat has started a program of analysis of IATTC observer data to look for instances of unusual patterns of reporting by observers. Advanced statistical analyses are being developed to detect observers whose data exhibit patterns of unusual set types given other data related to the set, unusually low dolphin mortality in circumstances in which higher mortality would be expected, and unusual patterns of tuna catches and dolphin mortality.

This work is at an early stage, but if it is successful, it will be more fruitful to apply this type of analysis to all observers than to focus solely on differences among programs.

APPENDIX A. DESCRIPTION OF THE RANDOMIZATION TEST

To statistically evaluate differences between programs, a randomization test was used to obtain an estimate of the probability that an average annual difference as large as, or larger than, that observed could be due to the chance assignment of trips to programs. The test was performed by randomly assigning trips from the pooled IRP data set for a particular country, by year, to two programs, and then computing the simulated average annual difference in the quantity of interest (e.g., average mortality per set) between programs for the random sample of trips. A total of 4,999 random samples of trips were simulated. The *p*-value for this test was computed as the proportion of simulated average annual differences with an absolute value as large as, or larger, than that actually observed. These calculations represent an approximation to a two-tailed test of the null hypothesis: no difference between programs.