



Sensitivity analysis of residency and site fidelity estimations to variations in sampling effort and individual catchability

Análisis de sensibilidad en estimaciones de residencia y fidelidad al sitio a variaciones en el esfuerzo y capturabilidad individual

Eduardo Morteo^{1,2✉}, Axayácatl Rocha-Olivares³ and Rodrigo Morteo⁴

¹Instituto de Ciencias Marinas y Pesquerías, Universidad Veracruzana, Calle Hidalgo #617, Col. Río Jamapa, CP 94290, Boca del Río, Veracruz, México.

²Laboratorio Veracruzano de Mamíferos Marinos, Acuario de Veracruz, A.C., Blvd. Manuel Ávila Camacho S/N, Playón de Hornos, CP 91700, Veracruz, Veracruz, México.

³Departamento de Oceanografía Biológica, Centro de Investigación Científica y de Educación Superior de Ensenada, Carretera Ensenada-Tijuana 3918, 22860 Ensenada, Baja California, México.

⁴Isotec de México, Av. Pedro Loyola #108, Fracc. Playa Ensenada, 22800 Ensenada, Baja California, México.

✉ emorteo@gmail.com

Abstract. Mark-recapture techniques are fundamental for assessing marine mammal population dynamics and individual temporal patterns. Since biases imposed by field conditions are generally unknown, we simulated variations in sampling effort (m) and maximum individual catchability (r_{max}) to analyze their effects on residency levels measured through the number of recaptures (occurrence, O), duration of stay (permanence, P), and average recurrence (periodicity, I) relative to a reference level of exhaustive daily sampling frequency. The number of recorded individuals (D_r) was also used to determine the performance of the simulations. Results for standardized ($\hat{\cdot}$) parameters showed that occurrences (O_s) were proportional to m and were not influenced by r_{max} . Individual permanence (P_s) and individual periodicity (I_s) were 8-49% and 3-11.74 times lower than expected, respectively, depending on m and r_{max} . Also, O_s , P_s , and I_s were not influenced by study duration, thus inter-study comparisons are feasible if m and r_{max} are similar. D_r was 68-92% ($r_{max} = 0.01$) and 1-8% ($r_{max} = 1.0$) lower than expected depending on m . Longer studies were more accurate but greater effort did not significantly increase D_r estimates. The use of bimonthly sampling frequencies ($m = 0.07$) was barely accurate and predictions for incomplete datasets were poor. Survey field data were also analyzed from 14 published studies on 4 dolphin species and compared to daily sampling frequencies; resulting values for O_s , P_s , and D_r were 62.4-93.3%, 11.6-66.4%, and 2.4-33.8% lower than expected, respectively; also I_s was 2.3-7.3 times lower than expected. The model produced D_r values that were similar to population estimates from empirical data, and bias was smaller than 15% in 87.5% of the cases, thus simulation accuracy was deemed acceptable.

Key words: population dynamics, occurrence, permanence, periodicity, simulations, dolphins.

Resumen. Las técnicas de marcado-recaptura son fundamentales para evaluar la dinámica poblacional de mamíferos marinos y sus patrones temporales individuales. Se simularon variaciones en el esfuerzo (m) y la máxima capturabilidad individual (r_{max}) para analizar los sesgos en la residencia individual, medida como el número de recapturas (presencia, O), la duración de la estancia (permanencia, P) y la recurrencia promedio (asiduidad, I), en relación con un muestreo exhaustivo de frecuencia diaria. También se analizó el número de individuos registrados (D_r) para determinar el desempeño de las simulaciones. Los resultados de los parámetros estandarizados ($\hat{\cdot}$) mostraron que la presencia (O_s) fue proporcional a m sin ser influida por r_{max} . La permanencia (P_s) y la asiduidad (I_s) fueron 8-49% y 3-11.74 veces más bajas que lo esperado, respectivamente, dependiendo de m y r_{max} . Los valores de O_s , P_s e I_s no cambiaron con la duración del muestreo y las comparaciones entre estudios son posibles siempre que m y r_{max} sean similares. El valor de D_r fue entre 68-92% ($r_{max} = 0.01$) y 1-8% ($r_{max} = 1.0$) menor de lo esperado dependiendo de m . Aunque los estudios más largos fueron más precisos, un mayor esfuerzo no incrementó significativamente el valor de D_r . Los muestreos quincenales fueron poco precisos y las predicciones para datos incompletos fueron poco confiables. También se analizaron los datos de campo de 14 estudios publicados de cuatro especies de delfines y se compararon contra muestreos diarios; los resultados de O_s , P_s y D_r fueron 62.4-93.3%, 11.6-66.4% y 2.4-33.8% más bajos de lo esperado, respectivamente; además, I_s fue 2.3-7.3 veces más baja que lo esperado. El modelo produjo valores para D_r que fueron similares a las estimaciones poblacionales de los estudios empíricos, con un sesgo menor al 15% en 87% de los casos, por lo que se considera que la precisión del modelo es aceptable.

Palabras clave: dinámica poblacional, presencia, permanencia, asiduidad, simulaciones, delfines

Introduction

Information on population size and individual movement patterns is considered fundamental for the effective management of species. In marine mammal populations, these attributes are commonly assessed through visual censuses and mark-recapture techniques (Buckland and Garthwaite, 1981), the latter being mainly based on patterns of natural markings and non-invasive methods such as photo-identification (Ballance, 1990; Wells and Scott, 1990).

Temporal distribution patterns for individual bottlenose dolphins (*Tursiops truncatus*) have been documented in many studies assessing their site fidelity (Defran and Weller, 1999) or residency (Ballance, 1990). Information on site fidelity helps to easily evaluate habitat use for particular individuals by recording their recurrence in previously occupied zones. Generally, a temporal dimension is not included in such assessments, since it is implied by the study duration. Alternatively, greater detail on habitat use can be obtained through a residency analysis, which is explicit on the time that individuals spend in a given geographic area (Wells and Scott, 1990). Consequently, calculations are delimited as natural (e.g., seasons, years) or arbitrary (e.g., short, medium, or long) periods (Ballance, 1990). Although there are several methods to evaluate residency (Simões-Lopes and Fabian, 1999; Möller et al., 2002; Rossi-Santos et al., 2007; Balmer et al., 2008; Morteo, 2011), all include at least one of the variables proposed by Ballance (1990) [Ballance variables henceforth]:

1), occurrence (O) is the number of recaptures of an individual over a given period:

$$O_i = \sum_{j=1}^k c_{i,j}$$

Where $c_{i,j}$ is a binary value indicating positive or negative capture (1 or 0) of individual i on the sampling date j , and k is the total number of sampled dates.

2), permanence (P) is the time over which an individual was recorded, determined by the difference between its first and last sighting:

$$P_i = \text{Max}(F_c) - \text{Min}(F_c)$$

For all $c=1$, where F is the sampling period (in days) for each individual i .

3), periodicity (I) is the recurrence of the individual, determined by the inverse of the average time between consecutive recaptures (days^{-1}). As defined here, a higher periodicity is reflected through a higher frequency of

recaptures. The latter is analogous to Ballance's third parameter, whose magnitude varies inversely with periodicity (*i.e.* higher values actually represent lower periodicity); hence, we suggest the use of our definition:

$$I_i = \frac{\sum_{j=1}^k c_j - 1}{\sum_{j=1}^k (F_{c_{j+1}} - F_{c_j})}$$

For all $c=1$, where F is the sampling period

On the other hand, Fisher's discovery curves (Fisher et al., 1943) are a graphical technique commonly used to estimate temporal dynamics and population size (Baker et al., 2006). The curve is sensitive to temporal patterns of new individuals (or species), recapture rates, and to frequency and sampling efficiency as well (Otis et al., 1978; Bebbler et al., 2007; Bearzi et al., 2008). Estimations of sampling errors associated with Fisher's curves are well known when assessing species diversity ("rarefaction curves", Bebbler et al., 2007) and the number of individuals in populations (Buckland and Garthwaite, 1981; Baker et al., 2006); however, the influence of such errors on residency and/or site fidelity estimations remains to be evaluated. In this study we used computer simulations to measure individual residency levels through Ballance's variables, and analyzed their sensitivity to variations in sampling effort and maximum catchability relative to the expected values under an exhaustive sampling design. Sampling effort was expressed as the survey frequency standardized by the duration of the study, and maximum catchability as the maximum probability of successfully identifying individuals, expressed as the maximum individual recapture rate for a given population. Also, Fisher's curves were used to determine the performance of the computational algorithm by comparing non-simulated empirical data from 14 published field surveys to the expected values from simulated exhaustive sampling frequencies.

Materials and methods

Simulations. A desktop application (Resident 2.10) was designed in Borland C++ Builder 6 to produce simulations of mark-recapture surveys (see Electronic Supplemental Material). Simulations are based on daily surveys, since most studies on population size and residency disregard repeated individual encounters during the same day (e.g., Fury and Harrison, 2008). The system uses 3 input parameters: 1), size of marked population (D); 2), survey period (t) (measured as the time in days between the first and last surveys), and 3),

maximum individual recapture rate for that period (r_{max}). Individuals are the basic unit of observation, thus the probability of capturing a previously marked or a new individual depends on each particular recapture rate (r_i). The system can either process actual empirical data or generate its own; in the latter case it allows the user to configure the number of random simulations (using the pseudo-random number generator from the system) based on input parameters. A run consists in computing individual values for Ballance's variables (O , P , and I), such that their statistical distribution is modulated by the input values, as well as Fisher's discovery curves for each data set (empirical or simulated). The output includes a summary of all the simulations in a format compatible with standard spreadsheets for further analyses.

The simulation algorithm assumes the ideal conditions discussed in Baker et al. (2006), Bebbler et al. (2007), and Bearzi et al. (2008), and also: 1), capture inequity: each individual is recognizable, and has a different probability of being marked and captured (r_i), the latter was assigned randomly and remained constant throughout the sampling period; 2), sampling independence: marking an individual does not affect its probability of being recaptured; 3), observer independence: all individuals keep their marks, and all observed individuals are recorded unmistakably; 4), geographic independence: all the study area has the same probability of being sampled, and individuals may be found anywhere; 5), social independence: the probability of marking an individual does not depend on the presence of others; 6), closed population in equilibrium: there is no immigration and the number of deaths equals births, and 7), binomial capture probabilities: a Bernoulli process determines whether individuals are observed (success probability) or not (failure probability) at a given time.

Experimental design. We considered 2 main sources of variation: 1), the sampling frequency (m), expressed as the number of sampling dates divided by the total time of the survey in days, assuming spatial and temporal homogeneity in sampling effort; therefore, the software randomly selected the sampling dates according to the survey design; and 2), random individual recapture rates (r_i), defined as the number of recaptures divided by the total number of sampling days (analogous to the estimation of "sightings per opportunity" from DeFran and Weller, 1999 or the "site fidelity index" from Simões-Lopez and Fabian, 1999).

Most photo-identification studies that have assessed residency or site fidelity have been applied to small cetaceans. Thus, we evaluated the experimental conditions of 14 empirical studies that have been carried out on 4 species of wild cetaceans published in peer-reviewed journals over the last 20 years. We focused only on those with asymptotic Fisher's curves, which suggested that most

of the marked individuals had been recorded by the end of the study (Table 1), such that D values would be closer to the actual size of the marked population. Thus it was assumed that these populations were virtually "closed" and also close to demographic equilibrium.

Different scenarios were simulated in order to quantify the bias introduced by sampling effort and individual catchability; we used the following input parameters based on actual field working conditions (Table 1): 1), the number of marked individuals (D) was set to 150; 2), study duration (t) lasted 3 years; 3), sampling frequencies (m) were monthly (0.03), bimonthly (0.07), 3 days per month (0.10), and weekly (0.14); for standardization purposes t and D are expressed as fractions. Finally, 4), maximum recapture rates (r_{max}) ranged within the following intervals delimited by a minimum of 2.74×10^{-3} (present only once a year) and 0.01, 0.10, 0.25, 0.50, 0.75, or 1.0; therefore different values for r_i were assigned randomly to each individual within the established maxima, using a uniform statistical distribution. The use of a uniform distribution is justified by the highly variable nature of r_i , which made fitting another probabilistic distribution difficult at best. We found the use of a uniform distribution a more suitable and parsimonious alternative, analogous to adopting a uniform prior in Bayesian computations (Jaynes, 2003).

We produced 10 000 simulated populations for each of the 6 intervals of r_{max} , and each population was sampled using the 4 frequencies of m (0.03, 0.07, 0.10, 0.14); an additional daily frequency ($m=1$) was simulated to represent the expected values under a more exhaustive sampling design. Each sampling date was considered a different experiment and analyses were carried out based on individual results (150 x 10 000 simulations, $N_i=150000$).

Residency and site fidelity. Residency was assessed through Ballance's parameters (O , P , and I) for all individuals in each of the 10 000 simulated populations for each experimental treatment. Treatments consisted in the combination of 6 maximum catchability levels (r_{max}), and 4 sampling conditions (m). In this study, site fidelity was considered analogous to Ballance's first variable (O), according to the definitions employed in most of the empirical studies (Table 1). Differences in Ballance's variables due to the experimental conditions (m and r_{max}) were standardized as fractions of the expected values (henceforth O_s , P_s , and I_s) based on daily samplings ($m=1$). For comparison with specialized literature, values for Ballance's variables in each experiment were averaged and then plotted along with their standard deviation. Differences among treatments were assessed with a non-parametric Kruskal-Wallis test ($\alpha=0.05$) with 4 blocks (m) and 6 levels (r_{max}). In addition, since many photo-

Table 1. Summary of empirical photo-identification studies in 4 species of wild dolphins. Variables on the left are sampling conditions: D = estimated population size; r_{max} = maximum recapture rate; t = study duration (in years); m = sampled days (fraction of t). Standardized (relative to a daily sampling frequency $m=1$) average percentage biases of residency variables are shown on the right: O_s = number of individual recaptures; P_s = time between first and last recaptures; I_s = inverse of the average time among consecutive recaptures; D_r = calculated deviation (%) from the simulated population size based on daily surveys. All values in parentheses show standard deviations

No.	Species	D	r_{max}	t	m	O_s	P_s	I_s	D_r	Reference
1	<i>Cephalorhynchus hectori</i>	22	0.51	2	0.11	89.2 (1.5)	20.6 (5.0)	19.3 (3.2)	4.3 (2.0)	Bejder and Dawson (2001)
2	<i>Tursiops truncatus</i>	155	0.17	1	0.27**	73.3 (2.8)	27.9 (7.6)	22.6 (1.6)	10.4 (2.4)	Ballance (1990)
3	<i>Tursiops truncatus</i>	40	0.67	1	0.15*	85.1 (2.1)	11.8 (4.2)	25.0 (3.3)	2.4 (1.6)	Williams et al. (1993)
4	<i>Tursiops truncatus</i>	82	0.28	2	0.16	83.6 (2.2)	18.7 (6.3)	23.6 (2.0)	6.8 (2.2)	Claridge (1994)
5	<i>Tursiops truncatus</i>	71	0.48	1	0.23**	76.9 (2.1)	11.6 (4.7)	24.0 (2.5)	4.1 (1.6)	Maize and Würsig (1999)
6	<i>Tursiops truncatus</i>	26	0.85*	2.4*	0.1	90.1 (1.4)	14.7 (4.3)	24.6 (8.0)	3.7 (1.6)	Simões-Lopez and Fabian (1999)
7	<i>Tursiops truncatus</i>	373	0.16	6	0.07	93.3 (1.7)	42.1 (9.5)	13.8 (1.5)	21.0 (4.5)	Defran and Weller (1999)
8	<i>Tursiops truncatus</i>	233	0.27	1	0.29**	70.7 (2.9)	17.5 (6.3)	25.3 (2.0)	6.8 (2.2)	Quintana-Rizzo and Wells (2001)
9	<i>Tursiops truncatus</i>	81	0.1	4	0.38**	62.4 (3.1)	30.9 (8.8)	21.2 (3.0)	13.8 (2.7)	Campbell et al. (2002)
10	<i>Tursiops truncatus</i>	313	0.04*	3	0.16*	83.9 (4.0)	66.4 (4.4)	43.3 (2.4)	33.8 (2.7)	Balmer et al. (2008)
11a	<i>Tursiops aduncus</i>	118	0.5	2	0.12	88.2 (1.8)	17.4 (4.8)	21.2 (3.0)	3.3 (1.6)	Möller et al. (2002)
11b	<i>Tursiops aduncus</i>	155	0.49	1.3	0.09	91.0 (1.5)	21.7 (4.9)	19.1 (3.2)	4.9 (2.0)	Möller et al. (2002)
12a	<i>Tursiops aduncus</i>	43	0.21*	3	0.08*	91.9 (2.3)	30.2 (8.7)	20.8 (1.3)	12.6 (1.7)	Fury and Harrison (2008)
12b	<i>Tursiops aduncus</i>	19	0.23*	3	0.09*	90.6 (1.9)	29.8 (6.7)	22.2 (1.0)	11.4 (2.4)	Fury and Harrison (2008)
13	<i>Sotalia guianensis</i>	58	0.1	3	0.24**	75.7 (2.4)	32.6 (8.8)	22.1 (1.3)	13.4 (2.7)	Rossi-Santos et al. (2007)
14	<i>Sotalia guianensis</i>	462	0.48	1	0.06	93.9 (1.3)	26.9 (6.3)	14.8 (1.9)	7.8 (2.1)	Espécie et al. (2010)
	Average	140.7	0.35	2.2	0.16	83.7	26.3	21.5	10.03	
	SD	135.8	0.19	1.46	0.09	9.3	13.6	5.3	8.1	

*= inferred value; **= outlier.

identification studies on small cetaceans last less than 3 years, we tested results for studies lasting 1 and 3 years for comparative purposes. Confidence intervals (95%) were also computed for each of the Ballance variables to assess the accuracy in the estimations.

Size of marked population. Fisher's discovery curves were computed for all simulated populations under the experimental treatments. Since D was fixed for all simulations, samplings were expected to achieve the total of marked individuals ($D=1$) within the study time (t) (Baker et al., 2006); therefore, the final number of recorded individuals (D_r) was measured as a fraction of D and examined through another Kruskal-Wallis test by blocks ($\alpha=0.05$). These analyses were also accomplished for simulations considering mark-recapture studies lasting 1 and 3 years, and confidence intervals (95%) were computed to assess the accuracy of the estimations.

Assessment of empirical sampling errors. Results from the previous sections allowed the assessment of sampling (m) and catchability (r_{max}) biases of the empirical studies (Table 1). Therefore, results from each paper were used as input values to run another 10 000 simulations in each case; we then evaluated the potential deviations relative to $m=1$ for each Ballance variable, expressing the deviations as fractions of the expected values (O_s , P_s , and I_s). Finally D_r values were compared to empirical estimations of population size (D) in each study to determine the performance of the simulation algorithm.

Results

Residency and site fidelity. Average standardized Ballance variables showed different trends; however, these were consistent when analyzed for studies lasting 1 and 3 years (variation was lower than 4.9% in all cases). Values for O_s changed significantly with m (Kruskal-Wallis, $p<0.05$), and were lower than expected (86-97% with respect to $m=1$) (Fig. 1). However, O_s did not show changes with r_{max} (Kruskal-Wallis, $p>0.05$) for a given m , and no interaction was found between variables (Fig. 1). It is also noteworthy that the occurrence was insufficient to compute the remaining Ballance variables when sampling conditions were the poorest (e.g., $m=0.03$ and $r_{max}=0.01$ or 0.1).

P_s values ranged between 8-49% lower than expected depending on m and r_{max} (Fig. 1); P_s was significantly lower (Kruskal-Wallis, $p<0.05$) for monthly surveys ($m=0.03$), while no significant differences were found when r_{max} was higher than 0.5 (Kruskal-Wallis, $p>0.5$). In this case there was interaction between variables ($p<0.05$).

I_s was between 3 to 11.7 times lower than expected depending on m and r_{max} (Fig. 1), and it was significantly lower only in the monthly surveys (Kruskal-Wallis, $p<$

0.05). Significant differences ($p < 0.05$) were found only for the most infrequent samplings and when r_{max} was lower than 0.5; in this case there was also interaction between variables ($p < 0.05$).

Confidence intervals showed the same pattern for all parameters (O_s , P_s , and I_s), in which accuracy was higher for the upper r_{max} (0.25-1.0) and m values (0.10 and 0.14), and decreased dramatically only for the lowest recapture rate ($r_{max} = 0.01$), such that in some cases no assessment could be made for P_s and I_s .

Size of marked population. As expected, average Fisher's curves showed that the number of recorded individuals (D_r) increases with m and r_{max} , and in general it reaches an asymptote when r_{max} is over 0.5. D_r values resulted in fractions of up to 68-92% ($r_{max} = 0.01$) and 1-8% ($r_{max} = 1.0$) of D for the most extreme values of m (Fig. 2).

Results showed significant differences among blocks (m) and treatments (r_{max}) (Kruskal-Wallis, $p < 0.05$) when the study lasted only 1 year. In this case, the number of marked individuals was significantly lower but only for monthly surveys ($m = 0.03$). On the other hand, D_r was significantly different ($p < 0.05$) for all r_{max} , and no interaction was found among variables ($p > 0.05$). A

three-year study also produced significant differences for the blocks (m) ($p < 0.05$) and treatments (r_{max}) ($p < 0.05$). Monthly surveys ($m = 0.03$) showed significantly lower values for D_r ($p < 0.05$) when r_{max} was lower than 0.5 ($p < 0.05$), therefore significant interactions occurred among variables ($p < 0.05$).

As expected, the curves that were closest to the expected population size ($D = 1$) were those that reached an asymptote over at least 10% of the study length (t) (Fig. 2). Also, D_r was the most distant from the expected values for the lowest r_{max} , thus being unreliable to estimate D . In general, the accuracy was higher for the higher r_{max} and tended to increase by the end of the surveys; conversely, accuracy was highest at $t = 0$ for the lowest r_{max} , and this was due to a lower range of capture possibilities that negatively biased the expected variability.

Assessment of empirical sampling errors. The 10 000 simulations based on input from empirical studies resulted in lower than expected values for O_s (62.4 to 93.3%) and P_s (11.6 to 66.4%) relative to a daily sampling frequency, where O_s was proportional to m (Table 1). Also I_s was between 2.3 and 7.3 times lower than expected (Table 1); according to our definition, the latter means that the interval

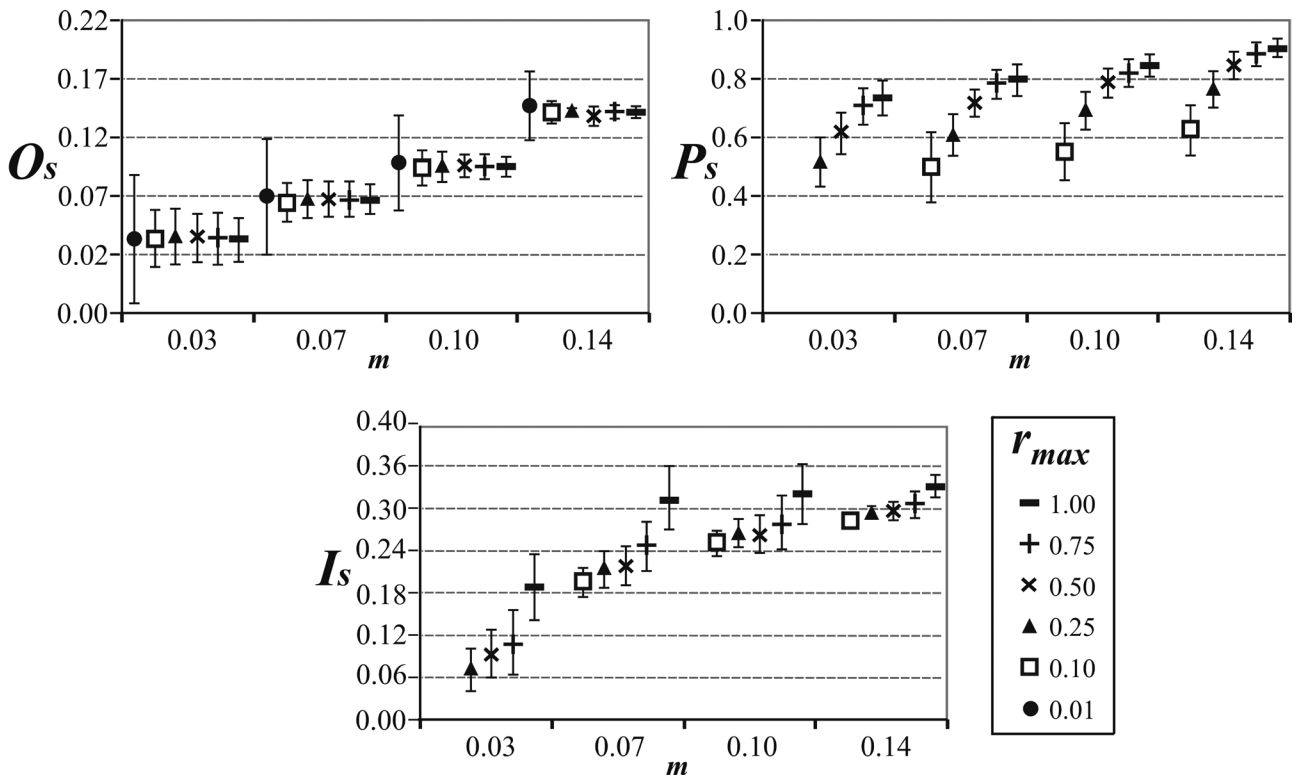


Figure 1. Average Ballance variables (\pm standard deviations). Occurrence (O_s = site fidelity), permanence (P_s) and periodicity (I_s) are standardized in relation to expected results based on an exhaustive (daily) sampling frequency ($m = 1$). Results were based on 10 000 simulations of closed populations under different m and maximum individual recapture rates (r_{max}).

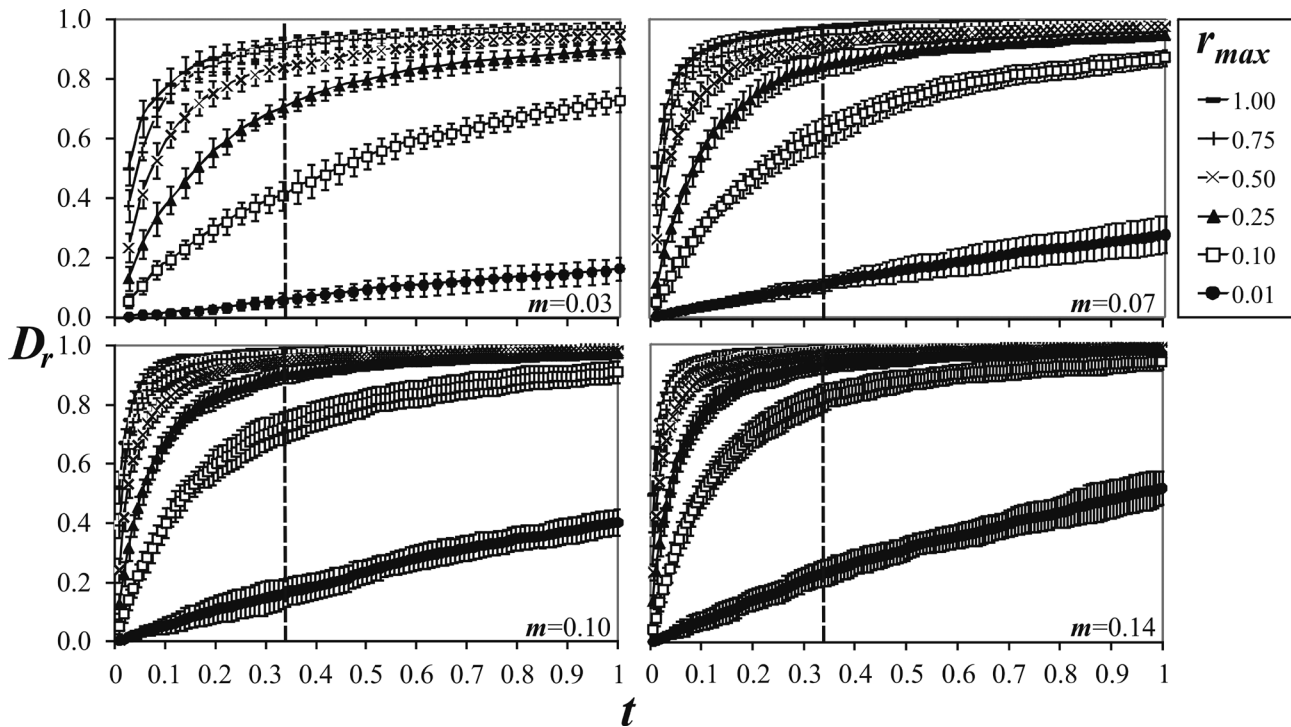


Figure 2. Average Fisher's individual discovery curves (\pm standard deviations) for 10 000 simulations of closed populations. The proportion of marked individuals (D_r) increases with time (t), for different sampling frequencies (m) and maximum individual recapture rates (r_{max}). The vertical dotted line establishes the final D_r assuming the end of a one-year study ($t = 0.33$).

between consecutive recaptures was actually larger, and by evaluating this variable the resulting residency is lower. Finally, D_r values differed between 2.4 and 33.8% with respect to the empirical D .

Discussion

Validity of model assumptions. Several authors have recognized the utility of models such as the one employed in this research for analyzing population dynamics of wild dolphins. Although results are still sensitive to assumption violations, most were met rationally. For instance, in most dolphin populations only a fraction of the individuals have conspicuous natural markings (Bearzi et al., 2008), and these may change with time; however, marked individuals may represent a considerable part of the population (e.g., 79% in Morteo, 2011), and most changes in markings are subtle, thus individuals may still be recognized if they are regularly photographed (Williams et al., 1993; Balmer et al., 2008).

Also, recapture probabilities may differ among population members due to age, sex, inconspicuous marks, low photographic quality, and avoidance of vessels, among other causes (Williams, et al., 1993; Simões-Lopes and Fabian, 1999; Balmer et al., 2008; Morteo, 2011).

Therefore, recapture heterogeneity was included *de facto* in the simulations using the entire possible range (Baker et al., 2006).

On the other hand, there may be concern about how well the algorithm simulates occurrences (O) in order to assign random individual recapture values (r_i). At least 2 studies have reported that the statistical distribution of O is not random (Quintana-Rizzo et al., 2001; Campbell et al., 2002). In this sense, despite the high variability in the statistical distribution of r , an earlier version of the software (Resident 1.3) showed that as sampling conditions worsened (e.g., $m < 0.1$ and $r_{max} < 0.25$) O 's probability distribution became negatively skewed, somewhat resembling results from actual field studies (Morteo, unpublished data). Such conditions were met in most simulations for the maximum recapture rates (r_{max}), but not for the sampling frequencies (m) (Table 1), thus there is a chance that O might be overestimated, especially in simulations with greater sampling effort. We argue that such an effect may be potentially insignificant since simulations for actual field data with the highest sampling frequencies showed the lowest bias, and the least variation in parameters (O_s , P_s , I_s , and D_r), thus being closest to the empirical results (Table 1).

The assumption of social independence is violated when individuals remain together for long periods; therefore, the

estimation of population sizes from the distribution of the number of recaptures can be highly biased (Darling and Morowitz, 1986). However, social bonds may be strong only in a few individuals (especially within *Tursiops* sp.) (Connor et al., 2000), and may develop within fusion-fission processes, allowing individual exchange among social units (Connor et al., 2000), thus such dependence may be weak for most of the population. For instance, Morteo (2011) concluded that bottlenose dolphins all around the world may associate with over 40 different dolphins during the course of a study, and many of these associations may occur by chance.

Finally, small cetacean populations are not necessarily closed or in demographic equilibrium; therefore, this method can only be applied if most of the marked individuals have been recorded, and encounters of new individuals have significantly decreased towards the end of the study. Such conditions were met in all cases, and constitute the basis for considering that a population is relatively closed at least during the sampling period.

Bearzi et al. (2008) concluded that most mark-recapture studies on small cetaceans may reasonably meet the model assumptions if the sampling period is short, individuals retain their marks, and survey effort is evenly distributed over the area. Therefore, we are confident that simulations in this study are valid and representative of many empirical field mark-recapture studies on wild dolphin populations. *Residency and site fidelity.* Sampling frequency (m) and recapture rates (r_i) are responsible for biases in the assessment of individual residency and site fidelity regardless of the duration of field surveys. However, given that sampling time represents a fraction of the total occurrence for each individual, the number of sightings (O) was proportional to the sampling frequency (m) multiplied by their recapture rate (r) (Fig. 1) (Defran and Weller, 1999). Since this parameter is analogous to the assessment of site fidelity, it may represent the most direct means for inter-study comparisons whenever sampling frequencies are comparable, regardless of the temporal patterns of individuals (Fig. 1).

The permanence of individuals (P) is not commonly used in the assessment of residency or site fidelity, given its direct dependence on the duration of field surveys. Finally, individual periodicity (I) is uncommon in residency assessment because it depends on both survey frequency and duration, thus being highly variable and presenting the greatest difficulties for inter-study comparisons.

It is remarkable that the confidence of the estimated parameters (O_s , P_s , I_s , and also D_r) was acceptable for most of the experimental conditions when recapture rates were above 10%; however, results become highly unstable and unreliable for the lowest r_{max} and the lowest m .

Population size. Several studies have shown that estimates of population size and the number of marked individuals are somehow positively correlated to survey effort and duration, but also to individual recapture rates (Baker et al., 2006; Bebber et al., 2007; Weir et al., 2008). This study showed that greater survey effort does not necessarily imply a significant increase in the number of marked individuals (Fig. 2). Also, even though longer studies approximated the population size (D) more accurately, we found no evidence that this affected the assessment of individual residency; therefore, a short study may provide a reasonable estimation of residency but confidence on it will strongly depend on effort intensity and extension.

Comparisons of simulated estimates for D_r (Table 1) resulted within the reported range of actual empirical estimates of population size (D) (studies 3, 4, 9, 11a, and 11b in Table 1); overestimations occurred only twice (studies 7 and 10) while underestimation occurred once (study 1). Thus the model proved to be robust given that such biases were lower than 10% in 9 instances (studies 1, 3, 4, 5, 6, 8, 11a, 11b, and 14) and were higher than 15% only twice (studies 7, 10). This reinforces the usefulness of Fisher's curves for estimating abundance in closed populations, as stated by Baker et al. (2006), but only when reliable assessments of uncertainty are available.

Nevertheless, our results also show that problems may arise while attempting to achieve predictions for incomplete datasets; this is either due to insufficient sampling duration, effort, or due to the high mobility of individual organisms (such as those in transit zones) (Fig. 2); therefore, unless observation time proceeds well beyond the asymptote (Baker et al., 2006), its usefulness for estimating D_r may be limited (Bebber et al., 2007).

Conclusions. The simulations showed that bimonthly sampling frequencies ($m < 0.07$) were barely enough to achieve acceptable estimates of O , P , I , and D_r , regardless of the duration of the study. Although estimations of D_r improved with greater effort, the change was insignificant for populations with low recapture rates ($r_{max} < 0.25$), yielding at least an underestimation of 30%. Since confidence intervals in estimations showed acceptable accuracy under most experimental conditions, except for the lowest r_{max} values, we argue that the assessment of residency is reliable. On the other hand, biases for the analyzed parameters (P and I) are significant, especially under low recapture rates ($r_{max} < 0.50$). We were also able to identify reference values that may help reduce data uncertainty, thus reinforcing the confidence and informative power of residency assessments. These are consistent with the classification of quality (Q) for capture-recapture data through closed population models as stated by Otis et al. (1978) ($Q =$ deficient, if $r_{max} < 0.05$; $Q =$ good, if $r_{max} > 0.35$).

Two main concerns relate to studies of residency and site fidelity: 1), no biases are accounted for sampling conditions, and 2), most studies use different definitions to assess 1 or both parameters. Methodological variations among studies are expected and this research provides alternatives to either assess the biases or standardize the results; however, there is a need to unify the criteria used for determination of residency for results to be directly comparable among studies (e.g., Möller et al., 2002; Fury and Harrison, 2008; Espécie et al., 2010). Therefore, we put forward 4 criteria to standardize residency determination: 1), individual recapture rates (r_i) must be higher than 25%; 2), Fisher's curve must be asymptotic for at least 10% of the study duration (e.g., Baker et al., 2006); 3), residency must be assessed over short periods, minimizing the possibilities of violating assumptions of closed and stable populations (e.g., migrations, seasonal and/or annual environmental oscillations), and 4), as most studies circumscribe their calculations in natural or arbitrary periods, the optimum number of residency categories (NRC) should be established empirically according to Balmer et al. (2008):

$$NRC = \frac{2 * IQR}{\sqrt[3]{D_r}}$$

where, IQR is the inter quartile range of the number of sightings, and D_r is the total number of animals sighted.

Acknowledgements

We thank Drs. Patricia Arceo, Luis Abarca, Héctor Perales, and Enriqueta Velarde from Universidad Veracruzana (UV) for their suggestions on early versions of this manuscript. Dr. Armando Jaramillo (National Institute of Ecology) helped to improve the theoretical and analytical framework for this research. We also acknowledge the valuable input from 2 anonymous reviewers, which substantially helped this manuscript to become more readable. This paper is part of the lead author's doctoral thesis at UV, Mexico, where he benefited from a doctoral fellowship from CONACyT. This work was supported by 2 PROMEP grants issued to the lead author (Apoyo a Nuevo PTC, and CAMyCRA), as well as the Marine Mammal Laboratory of Universidad Veracruzana and Acuario de Veracruz, A.C.

Literature cited

Baker, J. D., A. L. Harting and T. C. Johanos. 2006. Use of discovery curves to assess abundance of Hawaiian monk seals. *Marine Mammal Science* 22:847-861.

Ballance L. T. 1990. Residence patterns, group organization

and surface association of bottlenose dolphins in Kino Bay, Gulf of California, Mexico. *In* The bottlenose dolphin, S. Leatherwood and R. R. Reeves (eds.). Academic Press, San Diego, California. p. 267-284.

Balmer, B. C., R. S. Wells, S. M. Nowacek, D. P. Nowacek, L. H. Schwacke, W. A. McLellan, F. S. Scharf, T. K. Rowles, L. J. Hansen, T. R. Spradlin and D. A. Pabst. 2008. Seasonal abundance and distribution patterns of common bottlenose dolphins (*Tursiops truncatus*) near St. Joseph Bay, Florida, USA. *Journal of Cetacean Research and Management* 10:157-167.

Bearzi, G., S. Agazzi, S. Bonizzoni, M. Costa and A. Azzellino. 2008. Dolphins in a bottle: abundance, residency patterns and conservation of bottlenose dolphins *Tursiops truncatus* in the semi-closed eutrophic Amvrakikos Gulf, Greece. *Aquatic Conservation: Marine and Freshwater Ecosystems* 18:130-146.

Bebber, D. P., F. H. C. Marriott, K. J. Gaston, S. A. Harris and R. W. Scotland. 2007. Predicting unknown species numbers using discovery curves. *Proceedings of the Royal Society B: Biological Sciences* 274:1651-1658.

Bejder, L. and S. Dawson. 2001. Abundance, residency, and habitat utilization of Hector's dolphins (*Cephalorhynchus hectori*) in Porpoise Bay, New Zealand. *New Zealand Journal of Marine and Freshwater Research* 35:277-287.

Buckland, S. T. and P. H. Garthwaite. 1981. Quantifying precision of mark-recapture estimates using the bootstrap and related methods. *Biometrics* 47:255-268.

Campbell, G. S., B. A. Bilgre and R. H. Defran. 2002. Bottlenose dolphins (*Tursiops truncatus*) in Turneffe Atoll, Belize: occurrence, site fidelity, group size, and abundance. *Aquatic Mammals* 28:170-180.

Claridge, D. 1994. Photo-identification study to assess the population size of Atlantic bottlenose dolphins in central Abaco. *Bahamas Journal of Science* 5:12-16.

Connor, R. C., R. S. Wells, J. Mann and A. J. Read. 2000. *The bottlenose dolphin, Tursiops spp.: Social Relationships in a fission-fusion society. In* Cetacean societies: field studies of dolphins and whales, J. Mann, R. C. Connor, P. L. Tyack and H. Whitehead (eds.). University of Chicago Press, Chicago, Illinois. p. 91-126.

Darling, J. D. and H. Morowitz. 1986. Census of "Hawaiian" humpback whales (*Megaptera novaeangliae*) by individual identification. *Canadian Journal of Zoology* 64:105-111.

Defran, R. H. and D. W. Weller. 1999. Occurrence, distribution, site fidelity, and school size of bottlenose dolphins (*Tursiops truncatus*) off San Diego, California. *Marine Mammal Science* 15:366-380.

Espécie, M. A., R. H. O. Tardin and S. M. Simão. 2010. Degrees of residence of Guiana dolphins (*Sotalia guianensis*) in Ilha Grande Bay, south-eastern Brazil: a preliminary assessment. *Journal of the Marine Biological Association of the United*

- Kingdom 90:1633-1639.
- Fisher, R. A., A. S. Corbet and C. B. Williams. 1943. The relation between the number of species and the number of individuals in a random sample of an animal population. *Journal of Animal Ecology* 12:42-58.
- Fury, C. A. and P. L. Harrison. 2008. Abundance, site fidelity and range patterns of Indo-Pacific bottlenose dolphins (*Tursiops aduncus*) in two Australian subtropical estuaries. *Marine and Freshwater Research* 59:1015-1027.
- Jaynes, E. T. 2003. *Probability Theory: The Logic of Science*. Cambridge University Press.
- Maze, K. S. and B. Würsig. 1999. Bottlenose dolphins of San Luis Pass, Texas: Occurrence patterns, site-fidelity, and habitat use. *Aquatic Mammals* 25:91-103.
- Möller, L. M., S. J. Allen and R. G. Harcourt. 2002. Group characteristics, site fidelity and seasonal abundance of bottlenose dolphins *Tursiops aduncus* in Jervis Bay and Port Stephens, South-eastern Australia. *Australian Mammalogy* 24:11-21.
- Morteo, E. 2011. *Ecología social de delfines Tursiops truncatus en aguas costeras de Alvarado, Veracruz, México*. Tesis Doctoral, Instituto de Ciencias Marinas y Pesquerías, Universidad Veracruzana. Boca del Río. 129 p.
- Otis, D. L., K. P. Burnham, G. C. White and D. R. Anderson. 1978. Statistical inference from capture data on closed animal populations. *Wildlife Monographs* 62:1-135.
- Quintana-Rizzo, E. and R. S. Wells. 2001. Resighting and association patterns of bottlenose dolphins (*Tursiops truncatus*) in the Cedar Keys, Florida: insights into social organization. *Canadian Journal of Zoology* 79:447-56.
- Rossi-Santos, M. R., L. L. Wedekin and E. L. A. Monteiro-Filho. 2007. Residence and site fidelity of *Sotalia guianensis* in the Caravelas River Estuary, eastern Brazil. *Journal of the Marine Biological Association of the United Kingdom* 87:207-212.
- Simões-Lopes, P. C. and M. E. Fabian. 1999. Residence patterns and site fidelity in bottlenose dolphins, *Tursiops truncatus* (Montagu) (Cetacea, Delphinidae) off Southern Brazil. *Revista Brasileira de Zoologia* 16:1017-1024.
- Wells, R. S. and M. D. Scott. 1990. Estimating bottlenose dolphin population parameters from individual identification and capture-release techniques. *In Individual recognition of cetaceans: use of photo-identification and other techniques to estimation population parameters*, P. S. Hammond, S. A. Mizroch and G. P. Donovan (eds.). Special issue 12. International Whaling Commission, Cambridge, Massachusetts. p. 407-415.
- Williams, J. A., S. M. Dawson and E. Slooten. 1993. The abundance and distribution of bottlenose dolphins (*Tursiops truncatus*) in Doubtful Sound, New Zealand. *Canadian Journal of Zoology* 71:2080-2088.