ANALYSIS OF ANIMAL MOVEMENT USING OPPORTUNISTIC INDIVIDUAL IDENTIFICATIONS: APPLICATION TO SPERM WHALES

Hal Whitehead

Department of Biology, Dalhousie University, Halifax, Nova Scotia, Canada B3H 4J1

Abstract. Data sets in which animals are identified individually in different places and times may contain considerable information on movements. However, if the probability that an animal is reidentified depends on its movement pattern, then standard methods of analyzing movement are not applicable. I show that modifications of maximum likelihood methods, in which the identifications themselves establish the spatial and temporal distribution of effort, can be used to derive movement parameters in three situations: (1) Identifications in one location allow calculation of the "lagged identification rate" (the probability of reidentification after various time lags) as well as estimation of residence times inside, and outside, the study area. (2) When more than one study area is sampled, it is possible to derive lagged identification rates between them and to estimate movement rates between areas and other population parameters. (3) Movements through continuous space can be described by diffusion rates (rates of population spread), and plots of squared displacement against time lag. To simplify computation, and to permit the analysis of large data sets, summed nonindependent log-likelihoods can be maximized in place of the true log-likelihood to obtain approximately unbiased parameter estimates, and binomial, multinomial, or hypergeometric models can be approximated by the Poisson distribution. The first and third of the techniques were verified using simulated data, and all were applied to a 13-yr data set of identifications of sperm whales in the South Pacific Ocean. Residence times in waters close to the Galápagos Islands were of the order of 8 d, but during the study period there was a substantial net movement out of the Galápagos region and into waters of the coastal eastern tropical Pacific. Diffusion rates of sperm whales were ~700 km²/d over time scales from 1 to 100 d but decreased considerably over time scales of years, indicating displacements of ~50 km/d within home ranges spanning ~1000 km. Although giving relatively imprecise estimates of movement parameters compared to more standard methods, the techniques considered here should be particularly useful when examining animal movements over long time scales.

Key words: cetacean; diffusion; Galápagos Islands; jackknife; maximum likelihood; movement analysis; Physeter; *residence times; spatial scale; sperm whale; temporal scale.*

INTRODUCTION

The movement of animals through space is a key element of several areas of biology, including population biology (Andrewartha and Birch 1954), conservation biology (e.g., Soulé and Gilpin 1991), and behavioral ecology (e.g., Emlen 1991). Despite this, relative to many other areas of ecology, the study of movement has been less advanced, principally because of conceptual and technical difficulties (Nichols 1996, Turchin 1998). Turchin's (1998) recent book describes a range of techniques for analyzing movement within a conceptual framework based upon diffusion models. Diffusion describes how a population disperses through space, and the diffusion rate quantifies this. However, there are some situations for which well-developed analytical techniques are not available, and for which Turchin (1998) was only able to sketch a way forward. Such is the case for the analysis of movement using opportunistically collected individual-identification data.

When animals have been identified and reidentified (often analytically equivalent to "marked or tagged, and recaptured") over nonuniform and nonrandom spatial and temporal scales, so that the probability of recapturing an animal depends on its movement, then many of the standard methods of analyzing movement are not valid. In these situations, Turchin (1998) points to the powerful, and flexible, maximum likelihood technique of Hilborn (1990). In Hilborn's (1990) method, for any set of population and movement parameters and distribution of mark and recovery effort, the expected numbers of individuals with different resighting histories can be estimated using simulation, and compared with the observed numbers. Maximum likelihood estimation, using a Poisson model, is then employed to find the set of population and movement parameters which minimize the discrepancy between observed and expected resightings. Hilborn (1990) developed his model for a fisheries example with discrete areas in which fish were tagged and captured. However he notes that the general method is more broadly applicable.

Whales and dolphins (order Cetacea) operate over a particularly large range of spatial (≥8000 km; Stone et al. 1990) and temporal (≥100 yr; George et al. 1999) scales. For many years the subject of whale movements has been a key, and often contentious, issue in attempts to manage their populations (Donovan 1991). For instance, in their review of sperm whale (*Physeter macrocephalus*) stock structure, Dufault et al. (1999) conclude that ''to effectively conserve and manage sperm whales in the face of substantial anthropogenic disturbance, we need new and good information on modal and exceptional movement patterns over a range of time scales.'' Movements are potentially crucial elements of other areas of whale biology. For instance, movement patterns may have been significant in the evolution of cetacean social structures (Connor et al. 1998) and cultures (Rendell and Whitehead 2001).

An important potential source of information on whale movements comes from photographic identifications of individuals. For populations of right whales (*Eubalaena* spp.), humpback whales (*Megaptera novaeangliae*), fin whales (*Balaenoptera physalus*), blue whales (*Balaenoptera musculus*), killer whales (*Orcinus orca*), and sperm whales, photo-identification programs have been going on for at least a decade, cover large parts of the range of some populations, and include a substantial proportion of the animals in the populations (International Whaling Commission 1990). Photo-identification studies of smaller cetaceans, such as bottlenose dolphins (*Tursiops* spp.), also cover large time scales, but generally have less extensive spatial coverage (International Whaling Commission 1990, although see Defran et al. 1999). These data sets have been used to examine migratory destinations (e.g., Stone et al. 1990), to estimate population sizes using mark–recapture methods (e.g., Smith et al. 1999), and describe social organization (e.g., Christal et al. 1998). However, as most identifications include a position, these data sets also include a great deal of information on the movements of individuals over a range of spatial and temporal scales. This potential has not yet been properly tapped, partly because of the technical difficulties in analyzing movement patterns, and partly because of a lack of obvious analogs in methods developed for other species.

A particular problem with these data sets is that effort devoted to collecting individual identifications has been neither randomly nor systematically distributed in space and time. In most cases, identification records are highly clumped in certain ''study areas'' and ''field seasons,'' which may, or may not, correspond to changes in the distributions of the animals. A particular individual may have been recorded three times within a small study area over a one week period, not reidentified for another five years, and then perhaps observed in a location which was not being sampled originally. In these data sets there is usually no measure of effort other than the records of the identifications themselves.

My goal has been to find robust methods so that these data can be used to produce realistic models of whale movement which can then inform us about ecological and evolutionary questions. I will suggest how individual identification data can be used to estimate the parameters of three types of movement models, and to examine their legitimacy. The first type of model concerns movement into, and out of, a study area which occupies only a portion of the animals' range. This is extended to consider movement between study areas by making use of Hilborn's (1990) method, or other ''multistratum'' techniques (e.g., Brownie et al. 1993, Schwarz et al. 1993). Finally, movement through continuous space is considered, concentrating on estimating diffusion rates which indicate the rate at which animals disperse through space. In all three cases, I suggest alterations to the fairly standard techniques, which use maximum likelihood methods and a binomial–multinomial model, to make computation simpler and feasible on large data sets.

Methods of studying movement between study areas are discussed briefly, but, for the other two types of model, I suggest methods of examining trends in the data, maximum-likelihood techniques for estimating movement and population parameters, and test these techniques using simulated data. I then apply all three methods to photographic identifications of sperm whales in the South Pacific Ocean.

ANALYTICAL TECHNIQUES

Data sets

I assume that the data set consists of a series of sets of identifications of individual animals, together with the time and place of the identifications, so that the ith set took place at time t_i and in location $\mathbf{X}_i$ and was of n_i individuals. $\mathbf{X}$ can be a categorical variable giving study areas, or a one- to three-dimensional variable representing one- to three-dimensional space, m_{ij} represents the number of individuals identified in both set i and set j, τ_{ij} the time lag between the two identification sets ($\tau_{ij} = |t_j - t_i|$), and, if $\mathbf{X}$ is continuous, r_{ij} the displacement between them ($r_{ij} = |\mathbf{X}_i - \mathbf{X}_j|$). This, and other, notation is listed in Table 1.

Models of residency within, and emigration from, one study area

Here, I assume that all data are collected within a study area (so that X is single valued and redundant), and that we are interested in how animals move into, and out of, the study area. An intuitively reasonable manner of investigating this process is to examine the probability that an individual that was identified in the study area at time t would be identified during an identification of a randomly chosen resident of the study area at time $t + \tau$ later. I will call this probability the ''lagged identification rate'' of lag τ, $R(\tau)$ (following the use of ''lagged association rate'' in similarly struc-

TABLE 1. Principal notation.

Parameter/variable	Definition
a	mean residence time in study area before leaving
b	mean residence time outside study area before entering
f_j	probability that an animal in the study area on occasion j was identified
$g(\tau)$	numbers of pairs of identifications τ time units apart
i, j	used to denote a set of individual identifications collected at a particular time and location
m_{ij}	number of individuals identified in both set i and set j
$\hat{m}_{ij}$	expected number of individuals identified in both set i and set j
$m(\tau)$	number of reidentifications of the same individual τ time units apart
n_i	number of individuals identified in ith set
p	set of population/movement parameters
r	displacement (distance) through space
r_{ij}	displacement between identification sets i and j
$\{r_k\}$	displacements between successive identifications of the same individual
t_i	time of ith set of identifications
A_j	sampling area for occasion j
D	diffusion rate
$\hat{D}$	estimate of diffusion rate
L	log-likelihood or sum of (not necessarily independent) log-likelihoods
N	population size in a defined area
$P(\tau)$	probability that an individual in the study area is also in it after a lag of τ
$P(\tau, Y_1, Y_2)$	probability that an individual in study area Y_1 is in study area Y_2 after a lag of τ
$P(\tau, r)$	probability distribution function of moving a distance r during time τ
Q_{ij}	probability that an animal identified in set i is also identified in set j
$R(\tau)$	lagged identification rate, the probability an individual identified in a study area at time t is identified during a random identification in the study area at time $t + \tau$ later
$\hat{R}(\tau)$	estimate of lagged identification rate
$R(\tau, Y_1, Y_2)$	lagged identification rate between areas, the probability an individual identified in study area Y_1 at time t is identified during a random identification in study area Y_2 at time $t + \tau$ later
$\mathbf{X}_i$	location of ith set of identifications (can be categorical or one- to three-dimensional vector for continuous space)
Y_k	study area k
α	population density
δ	rate of mortality plus permanent emigration
λ	rate of emigration from a study area
μ	rate of immigration into a study area from wider population
ρ	rate of population increase
τ	time lag
τ_{ij}	time lag between identification sets i and j

tured analyses of social organization; Whitehead 1995). Calling $P(\tau)$ the probability that an individual in the study area at time 0 is also in it after a lag of τ, and N the number of individuals in the study area (initially assumed to be reasonably constant), then

$$R(\tau) = P(\tau)/N. \qquad (1)$$

The lagged identification rate, $R(\tau)$, can be obtained as a fitted function using observed data. A plot of $R(\tau)$ against lag τ, indicates how animals use the study area. The intercept on the y-axis is an estimate of $1/N$, and, if the lagged identification rate falls sharply at a lag of approximately $\tau = T$, then this indicates that the many animals are leaving the study area after residence periods of about this duration. A lagged identification rate falling to zero suggests permanent emigration or death, whereas a leveling off at large τ could be caused by some animals being permanent residents of the study area and/or others reimmigrating back into it. For any lag τ, the lagged identification rate, $R(\tau)$, can be estimated from the proportion of pairs of identifications τ time units apart which are of the same individual:

$$\hat{R}(\tau) = m(\tau)/g(\tau) \qquad (2)$$

where

$$m(\tau) = \sum_{i,j} \{m_{ij} | \tau_{ij} = \tau\} \qquad (3)$$

and $\qquad (4)$

$$g(\tau) = \sum_{i,j} \{n_i \times n_j | \tau_{ij} = \tau\}.$$

A plot of the lagged identification rate, $\hat{R}(\tau)$, against lag τ, calculated using Eqs. 2–4 is shown for a simulated data set in Fig. 1. Here it is apparent that many animals leave the study area after intervals of 1–5 d, but that either there are some permanent residents, or many return after leaving.

For many purposes it is desirable to fit a mathematical model to such residence data. Specifying a mathematical form and parameters for $P(\tau)$ is one way to formulate such a model. For instance, if the entire population is closed, and individuals within the study area migrate out of the study area with rate λ per unit
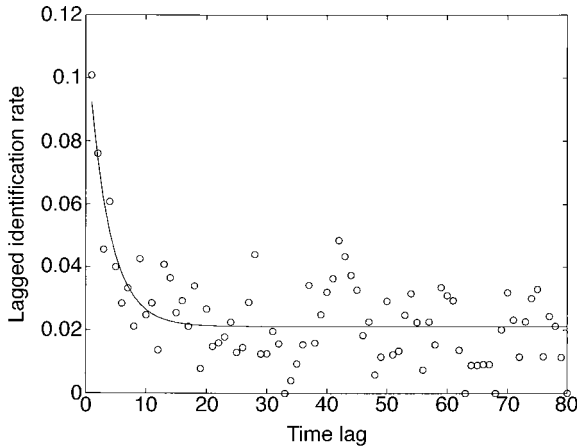
FIG. 1. Lagged identification rate against time lag (in arbitrary units) for one of the simulations described in Table 2, together with the expected lagged identification rate from a model of emigration and reimmigration fitted using maximum likelihood and binomial error (Eqs. 5 and 10).

time, and those outside it reimmigrate back into it with rate μ per unit time, a Markov process, then

$$P(\tau + \delta\tau) = P(\tau)(1 - \lambda\delta\tau) + [1 - P(\tau)]\mu\delta\tau$$

so

$$\delta P(\tau) = [\mu - (\lambda + \mu)P(\tau)]\delta\tau$$

and integrating,

$$[\log(\mu - [\lambda + \mu]P(\tau)) - \log(-\lambda)]/[-(\lambda + \mu)] = \tau.$$

Thus,

$$P(\tau) = (\lambda e^{-(\lambda+\mu)\tau} + \mu)/(\lambda + \mu). \quad (5)$$

If we are more interested in residence times than migration rates, then Eq. 5 can be rewritten as

$$P(\tau) = (be^{-(1/a+1/b)\tau} + a)/(a + b) \quad (6)$$

where a is the mean residence time within the study area, and b the mean residence time outside it. Such models can be extended in a number of ways. Mortality plus permanent emigration (at rate δ per time unit) can be added, so that Eq. 6 becomes

$$P(\tau) = e^{-\delta\tau}(\lambda e^{-(\lambda+\mu)\tau} + \mu)/(\lambda + \mu). \quad (7)$$

The parameters of these models can be estimated using mark–recapture methods in which the likelihood of the capture histories of all animals are calculated, and the parameters of the model (λ, μ, a, b, and/or δ in the above examples) are then chosen to maximize this likelihood (Sandland and Kirkwood 1981, Whitehead 1990). Unfortunately, if there are a large number of individuals and identification times as is sometimes the case in individual identification studies, the number of capture histories becomes very large and such computation is practically impossible (Whitehead 1990).

Here, I propose several simplifications to the mark–recapture methods so that the approach may be used on large data sets. These methods seem to give similar results to the mark–recapture methods on small data sets.

The probability that an animal identified on occasion i was also identified on occasion j is $P(\tau_{ij}) \times f_j$, where f_j is the probability that an animal in the study area on occasion j was identified. Then, using binomial theory, the likelihood that m_{ij} of the n_i animals identified on occasion i were also identified on occasion j is

$$I_{ij} = [P(\tau_{ij})f_j]^{m_{ij}} [1 - P(\tau_{ij})f_j]^{n_i-m_{ij}} \times \text{constant}. \quad (8)$$

Taking logs and summing over all pairs of occasions $\{(i, j)\}$, we obtain

$$L = \sum_i \sum_{j>i} [m_{ij}\log(P(\tau_{ij})f_j) + (n_i - m_{ij})\log(1 - P(\tau_{ij})f_j)]$$

$$+ \text{constant}. \quad (9)$$

Here L is not the log-likelihood of the entire data set, as the m_{ij}'s are not independent, rather it is the sum of a set of log-likelihoods, each with the same parameters and expected parameter values given by the log of Eq. 8. However, as shown analytically in the Appendix and by simulation later in the paper, the parameter estimates obtained by maximizing this sum are approximately unbiased. Thus useful estimates of the parameters determining $P(\tau_{ij})$ can be obtained by maximizing Eq. 9, although lack of independence means that standard likelihood methods should not be used to estimate confidence intervals or formally compare models. The summed log-likelihoods are used in place of the true log-likelihood because the latter becomes analytically intractable with many sampling periods.

Additional simplifications may be helpful. If there are more than a few sampling occasions, estimating all the f_j's using maximum likelihood is impracticable (Whitehead 1990). A useful simplification is to replace f_j by its unbiased estimator n_j/N (as in Seber 1982:557 and Whitehead 1990), so that Eq. 9 becomes

$$L = \sum_i \sum_{j>i} [m_{ij}\log(P(\tau_{ij})[n_j/N]) + (n_i - m_{ij})$$

$$\times \log(1 - P(\tau_{ij})[n_j/N])] + \text{constant}. \quad (10)$$

It is now only necessary to estimate N rather than all the f_j's.

This binomial model assumes that, on each sampling occasion, the effort directed to identifying each animal (f_j) was fixed in advance. In some situations it may be more appropriate to assume a fixed number of identifications (n_j). However, such an assumption leads to the less mathematically tractable hypergeometric distribution, and is unlikely to have much effect on the parameter estimates in most cases, as, unless m_{ij} is close to n_j, the hypergeometric and binomial distributions will be very similar.

If $n_j \ll N$, we can use the Poisson approximation to the binomial process, further simplifying the mathematics and computing (Hilborn 1990). In this case the summed log-likelihood is as follows:

TABLE 2. Estimation of parameters of a simple emigration–reimmigration model from simulated individual identification data.

Estimation method	Mean estimate (SD)		
	Emigration rate, λ	Immigration rate, μ	Population size, N
Theoretical	0.1	0.03	11.5
Poisson	0.1077 (0.0482)	0.0311 (0.0137)	11.59 (2.04)
Binomial	0.1077 (0.0468)	0.0313 (0.0138)	11.59 (1.98)
Mark–recapture	0.1026 (0.0167)	0.0305 (0.0050)	11.90 (1.20)

Notes: During each of 100 equal time intervals, members of a population of 50 individuals could move from a study area to surrounding areas (at a rate of $\lambda = 0.1$ individuals per time interval) or from surrounding areas into the study area (at a rate $\mu = 0.03$ individuals per time interval). Two hundred identifications were made by randomly choosing first a time interval, and then an individual within the study area during that interval. The mean number of individuals in the study area ($N = 50 \ \mu/[\lambda + \mu] = 11.5$), λ, and μ were estimated using the binomial model (Eq. 10), the Poisson model (Eq. 11), and a mark–recapture method (Whitehead 1990). The mean and SD of the estimates using each method are given for 20 runs of the model.

$$L = \sum_i \sum_{j>i} [m_{ij}\log(\hat{m}_{ij}) - \hat{m}_{ij}] + \text{constant} \quad (11)$$

where the expected number of reidentifications, $\hat{m}_{ij}$, is given by

$$\hat{m}_{ij} = n_i P(\tau_{ij}) f_j$$

or

$$\hat{m}_{ij} = P(\tau_{ij}) n_i (n_j/N) \quad \text{(making the same simplification as in the binomial model).} \quad (12)$$

The parameters of the residency model (e.g., λ, μ, a, b, δ), and number of animals in the study area, N, are then chosen to maximize the binomial (Eq. 10) or Poisson (Eq. 11) summed log-likelihood.

So far, I have assumed a constant population size in the study area, but an increasing or decreasing population can be modeled by using variable values of N, so by replacing N by $N(t_j)$ in Eqs. 10 (binomial model) or 12 (Poisson model). For instance, if ρ is the rate of population increase, to represent a population that is exponentially increasing ($\rho > 0$) or declining ($\rho < 0$), use

$$N(t) = N_0 e^{\rho t_j}. \quad (13)$$

In this case N_0 and ρ are estimated along with the residence parameters.

With large data sets, even using these simplifications, the number of comparisons made can be very substantial and computationally time consuming. Using the Poisson approximation helps, because Eq. 11 can be rewritten as

$$L = \sum_\tau [m(\tau)\log(P(\tau)) - P(\tau) \times g(\tau)/N]$$
$$+ \text{constant}. \quad (14)$$

Thus, the data are pooled by time lag τ, and each calculation of L is much cheaper. Another method of reducing computation is to restrict comparisons to those pairs of identifications less than a certain maximum time interval apart, T, so that the x-axis of the lagged identification rate plot extends up until lag T, and Eq. 14 only includes values of $\tau < T$.

Simulation of the simple model of reimmigration (Eq. 5) indicates that both binomial and Poisson approaches produce nearly unbiased estimates of the parameters λ, μ, and N but are somewhat less precise than the more complex, and much more computationally expensive, mark–recapture analysis (e.g., Table 2). Thus, at least in this situation, the approximations used to produce an efficient routine do not seem to be affecting parameter estimates in any important way (in the simulations the sampling was of a fixed number of identifications at each occasion, the scheme leading theoretically to the hypergeometric distribution). The procedure outlined above has the advantages over mark–recapture models which include residence (see also Seber 1982) of speed of computation (and thus possibility of computation for large data sets), simplicity, and that it can easily accommodate changes to the model (e.g., adding mortality is accomplished by making a small change to Eq. 5 to produce Eq. 7).

Models of movement between study areas

We can generalize the methods outlined above to investigate situations in which animals are identified in several different areas, so that $\{X_i\}$ are members of the set of study areas $\{Y_k\}$.

For instance, a lagged identification rate between different areas, $\hat{R}(\tau, Y_1, Y_2)$, can be estimated by adding a spatial component to Eq. 3 (and similarly to Eq. 4):

$$m(\tau, Y_1, Y_2)$$
$$= \sum_{i,j} \{m_{ij} | \mathbf{X}_i = Y_1, \mathbf{X}_j = Y_2, t_j - t_i = \tau\}. \quad (15)$$

The lagged identification rate between two sets of identifications now depends on the positions of the identifications, as well as the time lag.

Similarly, a spatial component can be added to the likelihood equations by modifying Eq. 9 so the summed log-likelihood becomes

$$L = \sum_i \sum_{j>i} [m_{ij}\log(P(\tau_{ij}, \mathbf{X}_i, \mathbf{X}_j) \times f_j) + (n_i - m_{ij})$$

$$\times \log(1 - P(\tau_{ij}, \mathbf{X}_i, \mathbf{X}_j) \times f_j)]$$

$$+ \text{ constant.} \tag{16}$$

Here, $P(\tau_{ij}, \mathbf{X}_i, \mathbf{X}_j)$ is the probability of moving from area $\mathbf{X}_i$ to $\mathbf{X}_j$ in time τ_{ij}, and is specified by generalizations to several areas of equations such as Eq. 5 (e.g., Brownie et al. 1993, Schwarz et al. 1993). Multinomial likelihood equations are used in the ''multistratum'' capture–recapture methods described by Brownie et al. (1993) and Schwarz et al. (1993). Unfortunately, principally because of the difficulties of calculating the $P(\tau_{ij}, \mathbf{X}_i, \mathbf{X}_j)$'s, such multinomial methods only work with a few ($<\sim 4$) study areas and a few ($<\sim 4$) sighting occasions (Brownie et al. 1993).

To get around this limitation, some of the same simplifications can be made as in the previous section with just one study area. First, f_j can be replaced by $n_j/N(\mathbf{X}_j)$:

$$L = \sum_i \sum_{j>i} [m_{ij}\log(P(\tau_{ij}, \mathbf{X}_i, \mathbf{X}_j) \times n_j/N(\mathbf{X}_j)) + (n_i - m_{ij})$$

$$\times \log(1 - P(\tau_{ij}, \mathbf{X}_i, \mathbf{X}_j) \times n_j/N(\mathbf{X}_j))]$$

$$+ \text{ constant.} \tag{17}$$

We can also use the Poisson approximation (justifiable if $n_j \ll N(\mathbf{X}_j)$, Hilborn 1990):

$$L(\mathbf{p}) = \sum_i \sum_{j>i} [m_{ij}\log(\hat{m}_{ij}(\mathbf{p})) - \hat{m}_{ij}(\mathbf{p})]$$

$$+ \text{ constant} \tag{18}$$

where

$$\hat{m}_{ij} = P(\tau_{ij}, \mathbf{X}_i, \mathbf{X}_j) \times n_i \times n_j/N(\mathbf{X}_j). \tag{19}$$

However, the problem of calculating the $P(\tau_{ij}, \mathbf{X}_i, \mathbf{X}_j)$'s remains. In situations with few (~ 2–3) study areas, recurrence equations can be used. Hilborn (1990) presents a powerful and general technique for circumventing such limitations. His method consists of first constructing a simulation population and movement model that describes how individuals move between areas. This model contains some unknown parameters (for instance the rates of mortality and movement between areas, and the population sizes in the different areas) here designated by the vector $\mathbf{p}$. Then, given $\{n_i\}$, $\{t_i\}$, and $\{\mathbf{X}_i\}$, and any set of parameters for the population-movement model ($\mathbf{p}$), the simulation model can be run a number of times to derive expected numbers of resightings between areas: $\{\hat{m}_{ij}(\mathbf{p})\}$. Estimates of the parameters in the population-movement model ($\mathbf{p}$) can be obtained by finding those that maximize the likelihood (Hilborn 1990) or (as justified in *Appendix*) the summed log-likelihoods, $L(\mathbf{p})$.

Hilborn (1990) discusses a variety of possible extensions to his technique, methods of finding confidence intervals for the parameters, and analysis of the fit of the model to the data using the method of de-

viances. Temporal changes in population sizes can also be added by making the $N(\mathbf{X}_j)$'s time dependent (i.e., $N(t_j, \mathbf{X}_j)$).

General models of movement in continuous space

While the approach of Hilborn (1990) is very powerful and flexible (Turchin 1998), it cannot be directly applied to a situation common in studies of whale populations using photo-identification: the identification of individuals opportunistically in continuous space and time. The study area could be divided into spatial strata, and movement probabilities between them used as parameters, but if this subdivision is coarse then much of the information on movements in the data set is lost, and if it is too fine then simulating population flow among so many strata is not feasible, even using the simplifications discussed above. However a method can be devised to deal with continuous space, if we make some simplifying assumptions.

In two dimensions, if mean population density is roughly uniform at all displacements from a randomly chosen location (at α animals/unit area), and the probability of identifying an animal in a given location is independent of the time since it was last identified (an assumption most obviously violated when tracking animals), then the summed log-likelihoods of the data under the binomial model is (as in Eq. 9)

$$L = \sum_i \sum_{j>i} [m_{ij}\log(Q_{ij}) + (n_i - m_{ij})\log(1 - Q_{ij})]$$

$$+ \text{ constant} \tag{20}$$

where Q_{ij} is the probability that an animal identified in set i is also identified in set j:

$$Q_{ij} = P(\tau_{ij}, r_{ij})\delta r(A_j/2\pi r_{ij}\delta r)f_j$$

$$= P(\tau_{ij}, r_{ij})A_j f_j/2\pi r_{ij}. \tag{21}$$

Here, A_j is the area sampled on occasion j, $P(\tau, r)$ is the probability density function for the displacement, r, over lag τ, so that the probability that, over lag τ, an animal has displaced between r and $r + \delta r$ units is $P(\tau, r)\delta r$. As previously, we can simplify, replacing f_j by $n_j/N(\mathbf{X}_j)$, and $N(\mathbf{X}_j)$, the population being sampled on occasion j, can be estimated by

$$N(\mathbf{X}_j) = (A_j\alpha + n_i P(\tau_{ij}, r_{ij})A_j/2\pi r_{ij}). \tag{22}$$

The term $n_i P(\tau_{ij}, r_{ij})A_j/2\pi r_{ij}$, the expected number of animals identified on occasion i in the area sampled on occasion j, is needed when τ_{ij} and r_{ij} are both small so that nearly the same set of animals are being resampled; the term becomes insignificant at larger τ_{ij} or r_{ij}. Replacing f_j by $n_j/N(\mathbf{X}_j)$ and using Eq. 22, Eq. 21 becomes

$$Q_{ij} = P(\tau_{ij}, r_{ij})A_j(n_j/2\pi r_{ij})/(A_j\alpha + n_i P(\tau_{ij}, r_{ij})A_j/2\pi r_{ij})$$

$$= P(\tau_{ij}, r_{ij})n_j/(2\pi r_{ij}\alpha + n_i P(\tau_{ij}, r_{ij})). \tag{23}$$

Then, the parameters that define $P(\tau, r)$, as well as α,

TABLE 3. Estimation of diffusion rate from simulated individual identification data using standard and maximum likelihood techniques with either binomial or Poisson error.

Study area	Figure panel	Estimated diffusion rate, D (SD) [theoretical $D = 0.01$]			Estimated density, α (SD) [theoretical $\alpha = 0.99$]		
		Standard	Binomial	Poisson	Calculated	Binomial	Poisson
All	a	0.0089 (0.0017)	0.0096 (0.0015)	0.0095 (0.0015)	1.02 (0.07)	0.87 (0.18)	0.86 (0.18)
Strip	b	0.0046 (0.0012)	0.0086 (0.0015)	0.0085 (0.0015)	1.00 (0.05)	1.02 (0.32)	1.02 (0.33)
Central square	b	0.0050 (0.0012)	0.0092 (0.0023)	0.0091 (0.0023)	0.99 (0.04)	0.90 (0.28)	0.91 (0.28)
Central rectangle	d	0.0037 (0.0013)	0.0093 (0.0035)	0.0091 (0.0034)	0.99 (0.06)	0.98 (0.43)	1.01 (0.43)
Bottom left, then top right	e	0.0064 (0.0012)	0.0098 (0.0023)	0.0097 (0.0024)	1.00 (0.06)	0.75 (0.25)	0.74 (0.26)
Top/bottom strips	f	0.0056 (0.0014)	0.0090 (0.0019)	0.0089 (0.0018)	0.99 (0.06)	0.66 (0.27)	0.62 (0.30)

Notes: One hundred individuals were placed randomly in a 10×10-unit habitat, and then each made 100 steps of length 0.2 units with the direction of each step being randomly chosen. Boundaries of the habitat were reflective. For six ''study areas'' (shown in Fig. 2), a random identification data set was compiled by randomly choosing an instantaneous time and a 1×1-unit square section of the study area, identifying all animals in the square at that time, and then repeating this process until 100 identifications were obtained. The diffusion rate, D, was estimated by the standard technique (Eq. 26), as well as by the binomial likelihood method (Eq. 20) and its Poisson approximation (Eq. 11). The mean density of animals as experienced by a randomly chosen animal (α animals per unit area; theoretically, $(100 - 1)/(10 \times 10) = 0.99$) was calculated by summing the number of animals in each 1×1-unit square at each time (''calculated'') and was also estimated using the binomial and Poisson likelihood methods. The mean and SD of the estimates for the 20 runs for each study area are shown.

can be estimated by maximizing Eq. 20, or its Poisson approximation, Eq. 11, with

$$\hat{m}_{ij} = Q_{ij} \times n_i. \tag{24}$$

One of the simplest forms of movement model is an uncorrelated random walk, in which the animal heads in a random direction at each of a sequence of very short steps. The uncorrelated random walk is well described by one parameter, the diffusion rate D, or rate of population spread through space (Turchin 1998). The diffusion rate can be formally defined in a number of equivalent ways (see Okubo 1975 or Turchin 1998). In the case of the uncorrelated random walk, for instance, it can be defined as the step length squared divided by four times the time between steps. The uncorrelated random walk also models, quite well, a number of situations where the movements of an animal are not actually random (Turchin 1998). For an uncorrelated random walk with diffusion rate D (from Turchin [1998], modifying Eq. A.15 for radial coordinates),

$$P(\tau, r) = \frac{r \exp\left(-\dfrac{r^2}{4D\tau}\right)}{2D\tau}. \tag{25}$$

The standard estimator for D in this situation is

$$\hat{D} = \sum_k r_k^2/4 \sum_k \tau_k \tag{26}$$

where the sums are over successive identifications of the same individual, the r_k's are the distances between these identifications, and τ_k's the time intervals between them (from Turchin 1998: Eq. 7.5).

Simulations suggest that the likelihood procedure outlined above (Eqs. 20–25), using either the binomial or Poisson model, gives approximately unbiased estimates of D (Table 3), even when the study area is irregularly and nonrandomly sampled. In these situa-

tions, the standard estimator of D (Eq. 26) performs poorly (Table 3; Turchin 1998:258).

It is customary, and often very useful, to examine movement patterns from individual mark–recapture data by plotting the mean-squared displacement over any time lag against the lag (Turchin 1998). For instance, a two-dimensional simple random walk predicts that this should be a linear function through the origin with the slope of the line being equal to four times the diffusion rate, D. However, if animals are not equally likely to be identified in different parts of their habitat, then this approach is flawed (see Fig. 3; $\times$'s lie below true line with all sampling schemes except case a, random identification throughout study area).

An alternative is to divide the time lags into intervals $\{0–\tau_1, \tau_1–\tau_2, \ldots\}$, and for each interval ($w = 1, \ldots$) estimate the diffusion rate, D_w, using the likelihood methods just described. Then the mean-squared displacement in each lag interval is estimated from

$$\hat{r}_w^2 = 4\hat{D}_w \times \text{mean}(\tau_{ij}| \tau_{w-1} < \tau_{ij} < \tau_w). \tag{27}$$

Simulations suggest that this method works satisfactorily, and better than the standard estimator, when effort is irregularly distributed (e.g., in Fig. 3, the maximum-likelihood estimators, $\bigcirc$ and $*$, are closer to the real data [dashed] line than is the standard estimator, $\times$). It is possible to estimate D and α separately for each time interval by maximizing the likelihood, or to estimate α for the entire data set, and then estimate D separately for each time interval, using the global estimate of α. The second procedure, using a global estimate of α, seems a little more accurate (e.g., in Fig. 3, the estimators using global estimates of α, $*$, are generally a little closer to the real data [dashed] line than those using separate estimates of α, $\bigcirc$).

If the data set is large, then it may make sense to pool ranges and time lags to shorten the calculations

and decrease use of computer memory. One method I have found useful is to use logarithmically increasing time bins to examine logarithmically increasing time lags.

Summary of method of estimation

Adopting the simpler, Poisson, distribution for errors, the three techniques outlined above can be broken down into the following steps:

1) Formulate a population and movement model with unknown parameters such as the rate of movement between study areas, mortality and the population density. These parameters are specified by the vector **p.**

2) Bin data by time and, perhaps, space, resulting in the numbers of animals identified in each bin (n_i), the mean time of identification (t_i), and the position $\mathbf{X}_i$.

3) Compare bins, calculating the number of animals in common (m_{ij}), time lag (τ_{ij}), and, if using continuous space, displacement (r_{ij}).

4) Construct a routine to calculate the expected numbers of reidentifications ($\hat{m}_{ij}$) between bins given the identification schedule $\{n_i, t_i, \mathbf{X}_i\}$ and population/movement parameters (**p**), and thus the log-likelihood or summed log-likelihoods, L, (from Eq. 11) either analytically, using recurrence relations, or by simulation.

5) Use an optimization routine to find those parameters (values of **p**) which maximize L.

Estimates of precision and hypotheses tests

The estimates of residence times, transition probabilities, diffusion rates, and other parameters produced using the methods described above have little utility without measures of precision. There are several possible approaches to estimating precision.

With just a few observation periods, it may be reasonable to compare only consecutive observation periods when calculating the m_{ij}'s, resulting in independent log-likelihoods, and so allowing standard likelihood methods to be used. The likelihood function in population/movement parameter space can be calculated and the ''support function'' (the difference between the log-likelihood at a particular combination of parameter values and the maximum log-likelihood) examined to indicate confidence regions (Edwards 1992). Alternatively, standard errors for parameters can be estimated from the inverse of the second partial derivative of the likelihood function (Silvey 1975).

These approaches assume independent data (m_{ij}'s), which will not be the case when summed log-likelihoods are used instead of the true log-likelihood. Alternatives for examining the precision of parameter estimates which can work with summed log-likelihoods include bootstrap (resampling the individuals with replacement to produce new data sets) or jackknife methods (creating new data sets with one individual omitted in each; Efron and Gong 1983).

If individuals themselves are not independent because they travel in fairly permanent groups, then boot-
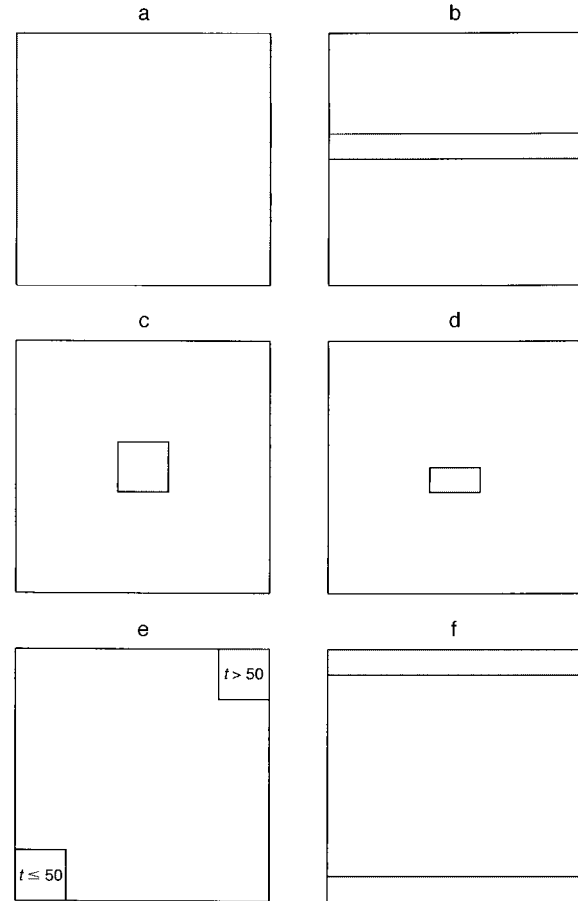

FIG. 2. Study areas (shaded) within a square habitat as used for the six sets of simulations described in Table 3.

strap or jackknife methods can be used with groups themselves being the units being included or omitted from each new data set. This is only possible if the compositions of the groups are known. If group membership is not known, but groups are suspected to exist, a procedure which seems to work, but has not been theoretically validated, is to jackknife on time periods rather than individuals, so that each jackknifed data set consists of the original data set missing those identifications collected during a time period. Standard errors of parameter estimates are calculated from these jackknifed estimates using the standard jackknife procedure (Efron and Gong 1983). Using this method on a simulated data set collected using the entire study area (Table 3, line 1; Fig. 2a) gave estimated standard errors of 0.0014 for D and of 0.15 for α, which can be compared with standard deviations from repeated simulated samples of 0.0015 and 0.18 respectively (Table 3).

If the elements of the calculation of likelihood can be considered independent then likelihood ratio tests (Silvey 1975) may be used to test the significance of adding parameters to a model (e.g., adding mortality, δ, to Eq. 5 to make Eq. 7). Even if not theoretically
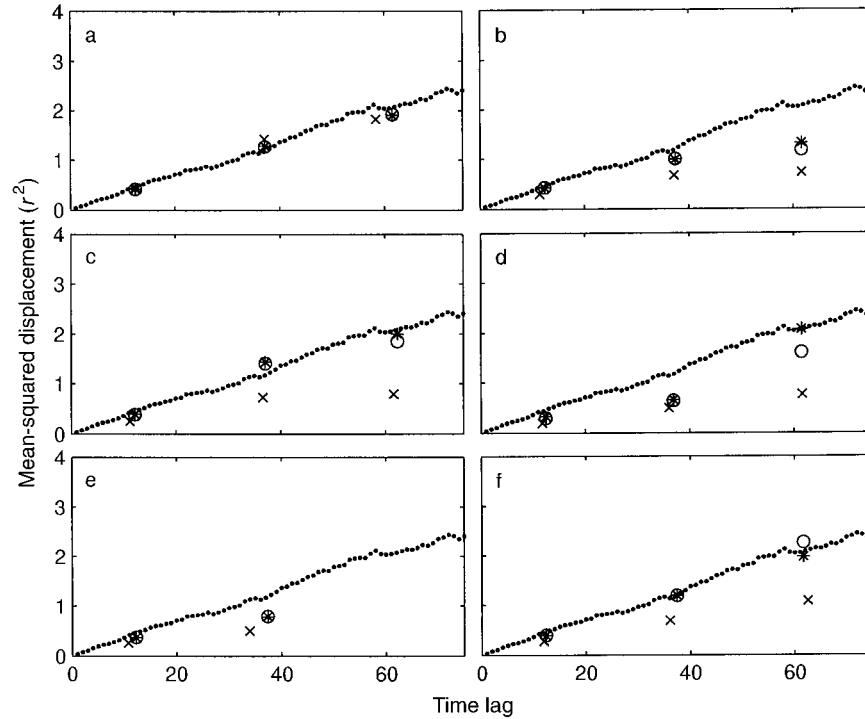
FIG. 3.   Mean-squared displacement as a function of time lag for simulations of the model described in Table 3. The true relationship between mean-squared displacement and time lag (averaged over all individuals and times) is shown by the dotted lines. For each of the six identification schemes, described in Table 3 and illustrated in Fig. 2 (except with 250 identifications rather than 100), and for three time intervals (1–25 units, 26–50 units, 51–75 units [51–75 units estimates unavailable for scheme e as there were no reidentifications over these time intervals]), the mean-squared displacement was estimated from the displacements between identifications of the same animal (standard method, $\times$), and from the estimated diffusion rate (using maximum likelihood) using Poisson error and estimates of density (α) calculated for each interval separately ($\bigcirc$) or globally ($\ast$).

TABLE 4.   Number of photographic identifications of individual sperm whales, and identified animals, during studies in the South Pacific Ocean, 1985–1997, in four different regions: Galápagos, coastal eastern tropical Pacific (coastal ETP), western Pacific, and southeastern Pacific (see Fig. 4).

Year	Months	Area	Identifications (accurate location $\leq$ 3 h)	Identified animals
1985	Feb	Coastal ETP	7 (7)	3
	Feb–Apr	Galápagos	633 (631)	282
1987	Jan–Jun	Galápagos	1095 (1052)	370
1988	Apr–Dec	Galápagos	346 (0)	136
1989	Jan–May	Galápagos	1350 (814)	569
1991	Jan–Mar	Coastal ETP	426 (426)	239
	Mar–Apr	Galápagos	113 (113)	72
1992	Jun	Coastal ETP	4 (4)	4
	Sep–Oct	Western Pacific	168 (167)	101
1993	Jan–Feb	Western Pacific	2 (1)	2
	Feb	Galápagos	18 (0)	16
	Feb–Apr	Southeast Pacific	201 (193)	110
	Apr	Coastal ETP	182 (182)	91
1994	Jan–Jun	Galápagos	32 (0)	22
1995	Apr–Jun	Galápagos	149 (149)	55
	Jun	Coastal ETP	13 (13)	9
1996	Apr–Jun	Galápagos	80 (0)	28
1997	Apr	Galápagos	146 (0)	24
Total			4965 (3752)	1851

PLATE 1. Fluke (tail) of a sperm whale showing marks used for individual identification off South Island, New Zealand. Photograph courtesy of L. Weilgart.

valid due to a lack of independence of data, a likelihood ratio test can indicate the significance of a particular parameter. For formal hypothesis tests in situations without independent data, then, custom permutation tests (Manly 1997) are perhaps the best option.

MOVEMENTS OF FEMALE AND IMMATURE SPERM WHALES

Sperm whales are large and sexually dimorphic (adult males, ~16 m long; adult females, ~11 m long) toothed whales (Rice 1989). Females and their young form groups of ~20 animals, and inhabit tropical and subtropical waters, whereas adult males are generally more solitary, and found at higher latitudes (Rice 1989). There is no indication that groups of females defend territories, and they range very widely (Best 1979, Whitehead and Waters 1990, Dufault and Whitehead 1995*b*).

Data set

The data set consists of 4965 high-quality (Arnbom's [1987] $Q \geq 4$) photographic individual identifications of 1851 individual female and immature sperm whales collected on 244 separate days from the waters of the South Pacific between 1985 and 1997 (Table 4, Fig. 4). Groups of sperm whales were tracked using passive acoustics for durations ranging from a few hours to a few days aboard 10–13 m long vessels (Whitehead and Gordon 1986). During these tracking periods, individuals were identified from photographs of the animals' flukes shown as they began long (~35 min), deep (~400 m) dives (Whitehead and Gordon 1986) (see Plate 1). Methods of taking and analyzing the photographs are described by Arnbom (1987) and Dufault and Whitehead (1995*a*). Each identification is linked to a date, and time (accurate within approximately ±5 min), and most are linked to a satellite-derived location. Locations were interpolated between fixes from a Tra-

cor Transtar SATNAV (Tracor, Austin, Texas, USA) during 1985–1991 (positions approximately every 1.5 h accurate to approximately ±0.5 km), and came from a Trimble Transpak GPS (Trimble, Sunnyvale, California, USA) during 1992–1995 (continuous positions accurate to approximately ±0.2 km).

The majority of the research was carried out off the Galápagos Islands (~0° N, 90° W), although identifications were collected right across the South Pacific, especially off the west coast of South America (Table 4, Fig. 4). Sperm whale abundance off the Galápagos Islands has fallen dramatically over the 13 yr of study, with many of the animals appearing to migrate to the waters off the South American mainland (Whitehead et al. 1997). The techniques described in the first part of this paper were used to investigate three general questions about movements of female and immature sperm whales: how long are they resident in waters close to the Galápagos Islands, at what rate have they moved from the general Galápagos region to waters off mainland South America, and how does the geographical displacement of a sperm whale group increase with time lag?

Because of the permanent groups formed by the whales and the fact that groups were tracked for periods of ~2 d, statistical methods which assume independence of individuals are inappropriate. Instead, to obtain standard errors for parameter estimates, I use the
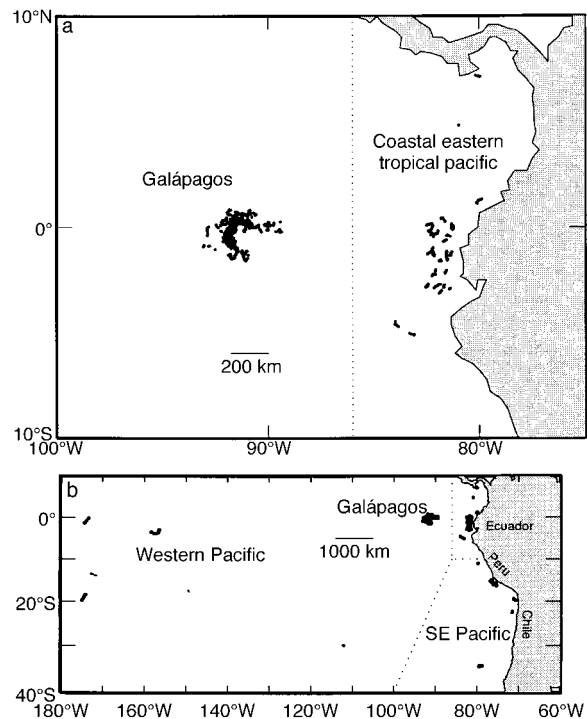

FIG. 4. Geographical distribution of photographic identifications of female and immature sperm whales in (a) the eastern tropical Pacific and (b) wider South Pacific, 1985–1997.
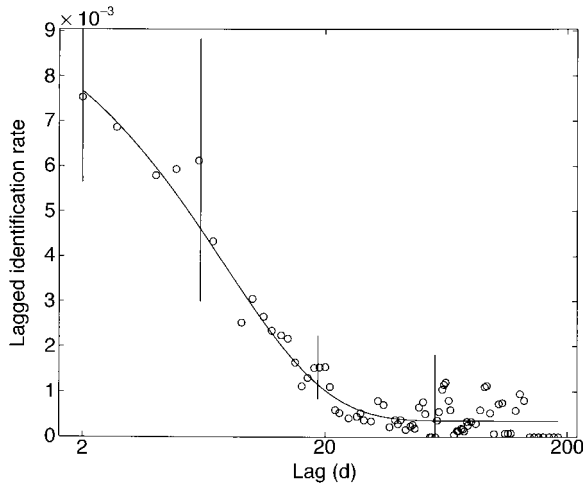
FIG. 5. Lagged identification rate for female and immature sperm whales in waters near the Galápagos Islands (○); bars show estimated standard errors, together with the expected lagged identification rate from a model of emigration and reimmigration (Eq. 5) fitted using maximum likelihood.

jackknife method, omitting consecutive 3–4 d periods in turn. Standard errors obtained in this way were little changed when jackknife periods of 30 d were substituted.

Residence in Galápagos waters

The data set was restricted to those photographs taken near the Galápagos Islands (Table 4) in order to examine residence in Galápagos waters (roughly defined by the distribution of effort shown on the left hand side of Fig. 4a). Lagged identification rates at lags from 2 to 200 d were calculated using Eq. 2. Lags of <2 d were not used as groups were often followed for two consecutive days (17 pairs of days). However, tracking of the same group rarely lasted more than 2 d (this happened on just three occasions).

The lagged identification rate is plotted against time lag in Fig. 5, smoothed by lumping lags into bins of median width of 4 d (bins narrower at shorter lags, and wider at longer lags). The lagged identification rate appears to fall precipitously at lags of ~3–10 d, settling at a fairly constant low level over longer lags.

The model of emigration from the study and reimmigration back into it (Eq. 6) was fit to the data using maximum likelihood with binomial loss. The model curve appears to fit the lagged identification rate well (Fig. 5), and has a substantially greater log-likelihood (−5556.01) than a simpler model in which emigration from the study area is deemed to be permanent, so that $b = \infty$ in Eq. 6, (log-likelihood −5691.66). Parameters of the model were estimated as follows (with estimated standard errors from the jackknife technique):

Mean residence in study area (a)

= 7.8 d (1 SE = 2.2 d)

Mean residence outside study area (b)

= 206.9 d (1 SE = 132.7 d)

Mean population in study area (N)

= 101 whales (1 SE = 37 whales).

This indicates that the Galápagos waters are a small part of the habitat of a population of several thousand whales, and that each group of whales spends of the order of 8 d near the islands at any time.

Movement from Galápagos region to the coastal eastern tropical Pacific

The movement of sperm whales between the general region of the Galápagos (which includes the smaller "Galápagos waters" study area used above) and the coastal eastern tropical Pacific, "coastal ETP," (Fig. 4a) is indicated by the lagged identification rate between the two areas (from Eqs. 2 and 15; Fig. 6). Only the rate from the Galápagos to the coastal ETP is shown, as data are too few for a useful plot for the reverse movements. In Fig. 6, it appears that individuals move between the two areas at rates of once every few years, as the lagged identification rate rises from zero for reidentifications within the same year to an approximately constant value after ~3.5 yr. These interarea movements were further investigated using the following model:

$N_G(t)$ = population in Galápagos region in year t,

$$t = 1985-1997$$

$N_E(t)$ = population in coastal ETP in year t

p_{GE} = probability of moving from Galápagos region to coastal ETP in any year

p_{EG} = probability of moving from coastal ETP to Galápagos region in any year

δ = annual mortality in both areas. Recruitment is assumed to be similar so that the overall population size does not change.

The expected values of the populations are given by the recurrence relations

$$\mathbb{E}(N_G(t+1)) = N_G(t) \times (1 - p_{GE}) + N_E(t) \times p_{EG} \quad (28)$$

$$\mathbb{E}(N_E(t+1)) = N_E(t) \times (1 - p_{EG}) + N_G(t) \times p_{GE}. \quad (29)$$

The probability that an animal identified in year y and area A (either Galápagos or coastal ETP) is present in year t and in each of the areas can be calculated from the recurrence relations:
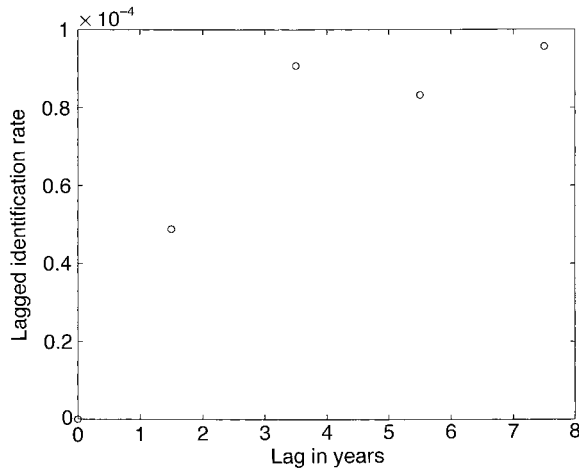
FIG. 6. Lagged identification rate for female and immature sperm whales from waters near the Galápagos Islands to the coastal eastern tropical Pacific.

$$P_{Ay}(G, t+1) = P_{Ay}(G, t)(1 - p_{GE})(1 - \delta)$$
$$+ P_{Ay}(E, t)p_{EG}(1 - \delta) \qquad (30)$$

$$P_{Ay}(E, t+1) = P_{Ay}(E, t)(1 - p_{EG})(1 - \delta)$$
$$+ P_{Ay}(G, t)p_{GE}(1 - \delta) \qquad (31)$$

and

$$P_{Ay}(G, y) = 1 \quad P_{Ay}(E, y) = 0 \qquad \text{if } A = G \quad (32)$$

$$P_{Ay}(E, y) = 1 \quad P_{Ay}(G, y) = 0 \qquad \text{if } A = E. \quad (33)$$

From these relationships, given the parameters $N_G(1985)$, $N_E(1985)$, p_{GE}, p_{EG}, and δ, the expected number of recoveries from area A to area B between years y and t can be calculated:

$$\hat{m}_{AyBt} = n_A(y)P_{Ay}(B, t)n_B(t)/N_B(t). \qquad (34)$$

From the observed and expected number of recoveries $\{m_{AyBt}, \hat{m}_{AyBt}\}$ the summed log-likelihood can be calculated from Eq. 11. Thus the parameters $N_G(1985)$, $N_E(1985)$, p_{GE}, p_{EG}, and δ were estimated by maximizing this likelihood. Using this procedure, the parameter estimators were:

Initial Galápagos population,

$$N_G(1985) = 4957 \text{ animals} \qquad (1 \text{ se} = 2184)$$

Initial coastal ETP population,

$$N_E(1985) = 4147 \text{ animals} \qquad (1 \text{ se} = 8171)$$

$$p_{GE} = 0.239/\text{yr} \qquad (1 \text{ se} = 0.112)$$

$$p_{EG} = 0.000/\text{yr} \qquad (1 \text{ se} = 0.001)$$

$$\delta = 0.021/\text{yr} \qquad (1 \text{ se} = 0.066).$$

Two simpler models were also fit to the data: a) $\delta = 0$ (i.e., no mortality or permanent emigration); b)

$N_G(1985) \, p_{GE} = N_E(1985) \, p_{EG}$ (i.e., migrations between the study areas are in equilibrium, and there is no net change in the population size in either area during the study period).

The log-likelihood of the full model was 456.74 (plus a large negative constant), that of model a was 456.48, and of model b 357.55. Theoretically, the difference in log-likelihood between models restricted by one parameter should be distributed as χ_1^2 if the more restricted model holds (Silvey 1975). Although likelihood ratio tests are theoretically invalid in this case, as the data are not independent, these differences in log-likelihood strongly suggest that mortality plus permanent emigration from both study areas is not significantly different from zero, but that there was a net migration from Galápagos to coastal ETP waters during the study period.

Using the estimated parameters shown above from the full model, the estimated changes in population size in the two study areas over the study period, as calculated from Eqs. 28 and 29, are shown in Fig. 7. The estimated change in the Galápagos population with time is very similar to that calculated using standard mark–recapture methods on nearly the same data set (Whitehead et al. 1997). However, this analysis quantifies the migration between the areas. Observed and expected reidentifications between years and study areas are given in Table 5. There are some substantial discrepancies, which were highlighted by an analysis of deviances. These can be attributed to the whales moving between areas, and being identified together, in permanent groups, so that movements and identifications are not independent. There was no obvious pattern to the deviances.

Sperm whale movement with time lag

For the analysis of displacement with time lag, I used all sperm whale identifications linked to a satellite-determined position within 3 h (Table 4). Because groups were followed over periods of 1–2 d, only lags
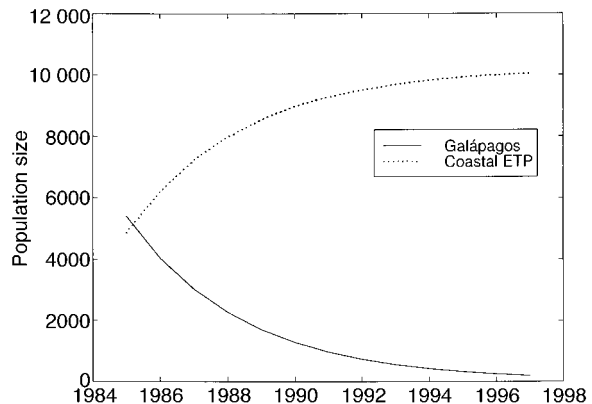

FIG. 7. Estimated changes in the numbers of female and immature sperm whales in the Galápagos region and coastal eastern tropical Pacific between 1985 and 1997.

TABLE 5. Observed and expected reidentifications of individual sperm whales between waters around Galápagos (G) and those in the coastal eastern tropical Pacific (10°N–10°S) (E) between 1985 and 1997.

Year and area	Year and area														
	1985 G	1986 E	1987 G	1988 G	1989 G	1991 G	1991 E	1992 E	1993 G	1993 E	1994 G	1995 G	1995 E	1996 G	1997 G
Observed reidentifications															
1985 G	282	0	49	0	29	10	0	0	1	1	0	2	0	0	0
1985 E	0	3	0	0	0	0	1	0	0	0	0	0	0	0	0
1987 G	49	0	370	0	56	13	0	0	2	8	9	13	0	18	0
1988 G	0	0	0	82	10	0	4	0	0	0	1	1	0	0	0
1989 G	29	0	56	10	569	21	7	0	0	12	4	6	2	4	0
1991 G	10	0	13	0	21	72	0	0	0	0	1	3	0	0	0
1991 E	0	1	0	4	7	0	239	0	0	0	0	0	0	0	0
1992 E	0	0	0	0	0	0	0	4	0	0	0	0	0	0	0
1993 G	1	0	2	0	0	0	0	0	16	0	3	0	0	0	0
1993 E	1	0	8	0	12	0	0	0	0	91	0	0	0	0	0
1994 G	0	0	9	1	4	1	0	0	3	0	22	12	0	10	0
1995 G	2	0	13	1	6	3	0	0	0	0	12	55	0	16	0
1995 E	0	0	0	0	2	0	0	0	0	0	0	0	9	0	0
1996 G	0	0	18	0	4	0	0	0	0	0	10	16	0	28	0
1997 G	0	0	0	0	0	0	0	0	0	0	0	0	0	0	24
Expected reidentifications															
1985 G	282	0	20	4	30	4	6	0	1	2	1	3	0	1	1
1985 E	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0
1987 G	20	0	370	10	70	9	7	0	2	3	2	6	0	3	2
1988 G	4	0	10	82	21	3	1	0	1	1	1	2	0	1	1
1989 G	30	0	70	21	569	24	7	0	5	4	7	17	0	8	7
1991 G	4	0	9	3	24	72	0	0	1	0	2	4	0	2	2
1991 E	6	0	7	1	7	0	239	0	0	2	0	0	0	0	0
1992 E	0	0	0	0	0	0	0	4	0	0	0	0	0	0	0
1993 G	1	0	2	1	5	1	0	0	16	0	1	2	0	1	1
1993 E	2	0	3	1	4	0	2	0	0	91	0	0	0	0	0
1994 G	1	0	2	1	7	2	0	0	1	0	22	3	0	1	1
1995 G	3	0	6	2	17	4	0	0	2	0	3	55	0	5	4
1995 E	0	0	0	0	0	0	0	0	0	0	0	0	9	0	0
1996 G	1	0	3	1	8	2	0	0	1	0	1	5	0	28	3
1997 G	1	0	2	1	7	2	0	0	1	0	1	4	0	3	24

Note: Sample sizes in each area and year are given along the diagonal.

of 3 d or more were considered. For each of seven time lag intervals, estimates of the diffusion rate, mean-squared displacement, and root-mean-squared displacement were calculated using the likelihood method described above with Poisson error and binning of time lag and displacement into sets of equal logarithmic size (factor of 2 for time lags, factor of 1.4 for displacements).

The results of the likelihood analysis are shown in Fig. 8. Over time lags from a few days to a few months, the movements of the sperm whales seem quite well described by a random walk with a diffusion rate of very approximately 750 km²/d (for lags of 3–100 d, estimated $D = 762$ km²/d, 1 SE = 495 km²/d). This compares with an estimate of 672 km²/d (1 SE = 141 km²/d) using the standard method of calculation (Eq. 26) over a lag of 1 d (Fig. 8). The standard method is valid over this short time lag as the whales were being tracked, and their probability of reidentification is unlikely to be much affected by their displacement. In contrast, over longer lags when whales moving greater distances are less likely to have been reidentified, the likelihood method is a more valid estimator. In contrast to the movement patterns over 1 to ~100 d, after in-

tervals of a year and more the rate of spread slows. The root-mean-squared displacement is plotted against time lag in Fig. 8c. This is less theoretically intuitive than the mean-squared displacement, but more easily interpretable as the approximate net distance traveled over different time scales. Over 3 d the female and immature sperm whales moved ~100 km, but during periods of years they move ~650 km, suggesting home ranges over multiyear time scales of the order of 1000 km across.

DISCUSSION

I hope that I have shown that a variety of aspects of movement can be studied using data sets in which animals are identified individually, but the identifications are distributed neither randomly nor uniformly in space or time and there is no independent measure of effort. Standard methods, in which it is assumed that animals are equally likely to be identified whatever their movement pattern, can produce very misleading results (e.g., Fig. 3, ×'s lie below true line with all nonrandom sampling schemes). By using maximum likelihood methods, and following the ideas of Hilborn (1990) and Turchin (1998), the identifications themselves can be
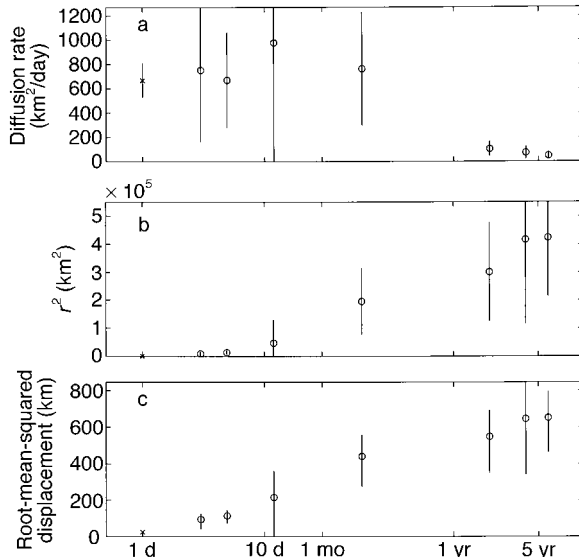
FIG. 8.   Results of likelihood movement analysis for female and immature sperm whales of the South Pacific: (a) estimated diffusion rates with time lag, (b) estimated mean-squared displacement with time lag, and (c) estimated root-mean-squared displacement with time lag. Vertical bars show estimated standard errors using the jackknife method, omitting three-day time periods in turn. Estimates are from likelihood estimation over longer time periods ($\bigcirc$), and the displacements between identifications of the same animal ($\times$) when tracking groups over short time periods.

used as a measure of effort, so that a variety of movement statistics can be estimated. These include residence times in one study area (e.g., Figs. 1 and 5), movement rates between study areas (Fig. 6), and diffusion rates (Figs. 3 and 8). Such information is vital for the investigation of many questions in population ecology, conservation, management, and evolution (Turchin 1998).

I have used modified maximum likelihood methods to investigate the movements of South Pacific sperm whales, indicating how animals use a small study area (Fig. 5), have redistributed themselves over scales of ~1000 km (Figs. 6 and 7), and move through the ocean over a range of time scales (Fig. 8). Over time scales of 1–100 d, sperm whale movements seem to be reasonably well described by an uncorrelated random walk, with diffusion rate of ~700 km²/d, and thus a root-mean-squared displacement of ~50 km/d. However, these movement rates cannot be extrapolated to longer time scales (Fig. 8), presumably because the animals have preferred ranges spanning ~1000 km within which they tend to stay.

These results have implications for the conservation and management of this species. For instance, intense shore-whaling from Paita, northwestern Peru (5° S, 81° W), removed ~7500 female sperm whales from the coastal eastern tropical Pacific waters between 1957 and 1981 (Clarke et al. 1980, Ramirez 1989). The movement analysis shows that this exploitation was

affecting whales over a much wider area than the whalers' operations, and was likely depopulating a large area of the eastern Pacific, despite a fairly stable catch per unit effort (Ramirez 1989). The movement analysis is also consistent with the suggestion that the decline in the Galápagos population during the late 1980s and early 1990s is due to animals moving into relatively productive coastal waters emptied by the Peruvian whaling, and not returning (Whitehead et al. 1997). By this analysis, I have begun to satisfy Dufault et al.'s (1999) call for information on modal movements of sperm whales over a range of time scales. Some of the results of this movement analysis have been used as input parameters for a model of the effects of exploitation on mobile sperm whale populations (Whitehead 2000), as well as during a consideration of the significance of movement in the evolution of cetacean culture (Rendell and Whitehead 2001).

The techniques illustrated in this paper can potentially be applied to a wide range of data sets in which moving organisms are identified individually by some means, and are then reidentified opportunistically. Identification can be photographic, genetic, or by means of artificial marks. Prime candidates are other cetacean photo-identification data sets (some prominent examples are given in *Introduction*). Artificial tagging data, in which "Discovery tags," stainless steel cylinders bearing serial numbers, were shot into whales and recovered during whaling operations (Brown 1978), have been used to examine movements in fairly simplistic ways (e.g., Best 1979). These data are more congruent to the fisheries tagging statistics used by Hilborn (1990) but perhaps could also be used to examine how displacement increases with time lag. Analysis of the movements of land birds using mark–recovery (usually leg-band) data is quite well developed (Nichols 1996). However, because bird movements can often be characterized by long-distance seasonal migrations between quite small home ranges, the questions asked by such studies (e.g., Nichols 1996) are rather different from those pertinent for more nomadic animals like oceanic cetaceans.

Although the likelihood methods developed in this paper seem useful, they are not without their limitations. Several approximations have been made in order to make the estimation computationally tractable: replacing the sampling rates, f_j, by n_j/N rather than estimating them using likelihood equations, maximizing the summed log-likelihoods rather than the true log-likelihood of the data, and using the Poisson model as an approximation to more theoretically valid binomial, multinomial, or hypergeometric models. Theoretical considerations (e.g., see Appendix; see also Hilborn 1990) and simulations (e.g., Tables 2 and 3) suggest that these simplifications generally do not bias parameter estimates in any important way. However, these issues should be borne in mind and tested using simulations as new applications are explored (Hilborn 1990). These methods should not

be used when the probability of identifying an animal in a particular location and time is dependent on its identification history (e.g., when tracking animals). I have found that the likelihood method of estimating diffusion rates is very sensitive to substantial recording errors, for instance if time or location is misreported so that an individual appears to have moved a very large distance in a short time, then improbable estimates of diffusion rates are obtained. But perhaps the most important drawback of these likelihood methods is imprecision. Where both standard and likelihood methods are valid, the likelihood methods are less precise (Tables 2 and 3, Fig. 8). This is presumably because the sighting data are being used to estimate effort as well as to define movements, while uniform or random distributions of effort are usually assumed in the standard methods. Substantial data sets are necessary to obtain reasonably precise estimates using the likelihood method. In this respect they seem less useful than movement analyses of continuous tracking data (e.g., Turchin 1998: chapter 5). However, animals can rarely be tracked, either physically or using radio tags, for more than a few months. It is also rare that identification schemes can be maintained so that the entire potential range of animal movement is sampled uniformly or randomly over long time scales. Therefore, I suspect that the likelihood methods developed here will be especially important over long (multiyear) time scales.

There are many ways that the techniques that I have outlined could be developed. For instance, correlated random walk models might be used to examine consistency of movement over shorter time scales, and variation between animals and classes of animals could be incorporated into the analyses.

There is no computer package which can perform all the analyses described in this paper. However, packages for the analysis of mark–recapture data (e.g., MARK; White and Burnham 1999) can be used to estimate transition rates between "strata", which can be geographical areas. However, because they use multinomial models, and for other reasons, these programs are only useful with small numbers of areas and identification occasions. Hilborn (1990) outlines how his simulation method, which can work with much larger data sets, may be programmed, relatively simply, to estimate movement rates between geographical areas. As an adjunct to a computer package for the analysis of social structure using individual identifications in the language MATLAB, I have written programs which are designed to make the first and third of the methods described in this paper easily applicable to a wide range of data sets (see *Supplementary Material*).

Literature Cited

Andrewartha, H. G., and L. C. Birch. 1954. The distribution and abundance of animals. University of Chicago Press, Chicago, Illinois, USA.

Arnbom, T. 1987. Individual identification of sperm whales. Reports of the International Whaling Commission **37**:201–204.

Best, P. B. 1979. Social organization in sperm whales, *Physeter macrocephalus*. Pages 227–289 *in* H. E. Winn and B. L. Olla, editors. Behavior of marine animals. Volume 3. Plenum, New York, New York, USA.

Brown, S. G. 1978. Whale marking techniques. Pages 71–80 *in* B. Stonehouse, editor. Animal marking. Macmillan, London, UK.

Brownie, C., J. E. Hines, J. D. Nichols, K. H. Pollock, and J. B. Hestbeck. 1993. Capture–recapture studies for multiple strata including non-Markovian transitions. Biometrics **49**:1173–1187.

Christal, J., H. Whitehead, and E. Lettevall. 1998. Sperm whale social units: variation and change. Canadian Journal of Zoology **76**:1431–1440.

Clarke, R., A. Aguayo, and O. Paliza. 1980. Pregnancy rates of sperm whales in the southeast Pacific between 1959 and 1962 and a comparison with those from Paita, Peru between 1975 and 1977. Reports of the International Whaling Commission (special issue) **2**:151–158.

Connor, R. C., J. Mann, P. L. Tyack, and H. Whitehead. 1998. Social evolution in toothed whales. Trends in Ecology and Evolution **13**:228–232.

Defran, R. H., D. W. Weller, D. L. Kelly, and M. A. Espinosa. 1999. Range characteristics of Pacific coast bottlenose dolphins (*Tursiops truncatus*) in the southern California Bight. Marine Mammal Science **15**:381–393.

Donovan, G. P. 1991. A review of IWC stock boundaries. Reports of the International Whaling Commission (special issue) **13**:39–68.

Dufault, S., and H. Whitehead. 1995a. An assessment of changes with time in the marking patterns used for photo-identification of individual sperm whales, *Physeter macrocephalus*. Marine Mammal Science **11**:335–343.

Dufault, S., and H. Whitehead. 1995b. The geographic stock structure of female and immature sperm whales in the South Pacific. Reports of the International Whaling Commission **45**:401–405.

Dufault, S., H. Whitehead, and M. Dillon. 1999. An examination of the current knowledge on the stock structure of sperm whales (*Physeter macrocephalus*) worldwide. Journal of Cetacean Research and Management **1**:1–10.

Edwards, A. W. F. 1992. Likelihood. Second edition. John Hopkins University Press, Baltimore, Maryland, USA.

Efron, B., and G. Gong. 1983. A leisurely look at the bootstrap, the jackknife, and cross-validation. The American Statistician **37**:36–48.

Emlen, S. T. 1991. Evolution of cooperative breeding in birds and mammals. Pages 301–337 *in* J. R. Krebs and N. B. Davies, editors. Behavioural ecology. An evolutionary approach, Third edition. Blackwell, Oxford, UK.

George, J. C., J. Bada, J. Zeh, L. Scott, S. E. Brown, T.

O'Hara, and R. Suydam. 1999. Age and growth estimates of bowhead whales (*Balaena mysticetus*) via aspartic acid racemization. Canadian Journal of Zoology **77**:571–580.

Hilborn, R. 1990. Determination of fish movement patterns from tag recoveries using maximum likelihood estimators. Canadian Journal of Fisheries and Aquatic Sciences **47**: 635–643.

International Whaling Commission. 1990. Report of the workshop on individual recognition and the estimation of cetacean population parameters, La Jolla, 1–4 May 1988. Reports of the International Whaling Commission (special issue) **12**:3–40.

Manly, B. F. J. 1997. Randomization, bootstrap and Monte Carlo methods in biology, Second edition. Chapman and Hall, London, UK.

Nichols, J. D. 1996. Sources of variation in migratory movements of animal populations: statistical inference and a selective review of empirical results for birds. Pages 147–197 *in* O. E. Rhodes, R. K. Chesser, and M. H. Smith, editors. Population dynamics in ecological space and time. University of Chicago Press, Chicago, Illinois, USA.

Okubo, A. 1975. Diffusion and ecological problems: mathematical models. Springer-Verlag, Berlin, Germany.

Ramirez, P. 1989. Captura de cachalote en Paita: 1976–1981. Boletin de Lima **63**:81–88.

Rendell, L., and H. Whitehead. 2001. Culture in whales and dolphins. The Behavioral and Brain Sciences **24(2)**:xx–xx.

Rice, D. W. 1989. Sperm whale. *Physeter macrocephalus* Linnaeus, 1758. Pages 177–233 *in* S. H. Ridgway and R. Harrison, editors. Handbook of marine mammals. Volume 4. Academic Press, London, UK.

Sandland, R. L., and G. P. Kirkwood. 1981. Estimation of survival in marked populations with possibly dependent sighting probabilities. Biometrika **68**:531–541.

Schwarz, C. J., J. F. Schweigert, and A. N. Arnason. 1993. Estimating migration rates using tag-recovery data. Biometrics **49**:177–193.

Seber, G. A. F. 1982. The estimation of animal abundance and related parameters. Second edition. Griffin, London, UK.

Silvey, S. D. 1975. Statistical inference. Chapman and Hall, London, UK.

Smith, T. D., J. Allen, P. J. Clapham, P. S. Hammond, S. Katona, F. Larsen, J. Lien, D. Mattila, Palsbøll, J. Sigurjónsson, P. T. Stevick, and N. Øien. 1999. An ocean-basin-wide mark–recapture study of North Atlantic humpback whales (*Megaptera novaeangliae*). Marine Mammal Science **15**:1–32.

Soulé, M. E., and M. E. Gilpin. 1991. The theory of wildlife corridor capability. Pages 3–8 *in* D. A. Saunders and R. J. Hobbs, editors. Nature conservation. 2. The role of corridors. Surrey Beatty, Chipping Norton, UK.

Stone, G. S., L. Flórez-Gonzalez, and S. Katona. 1990. Whale migration record. Nature **346**:705.

Turchin, P. 1998. Quantitative analysis of movement. Sinauer Associates, Sunderland, Massachusetts, USA.

White, G. C., and K. P. Burnham. 1999. Program MARK: survival estimation from populations of marked animals. Bird Study (Supplement) **46**:120–138.

Whitehead, H. 1990. Mark–recapture estimates with emigration and re-immigration. Biometrics **46**:473–479.

Whitehead, H. 1995. Investigating structure and temporal scale in social organizations using identified individuals. Behavioral Ecology **6**:199–208.

Whitehead, H. 2000. Density-dependent habitat selection and the modeling of sperm whale (*Physeter macrocephalus*) exploitation. Canadian Journal of Fisheries and Aquatic Sciences **57**:223–230.

Whitehead, H., J. Christal, and S. Dufault. 1997. Past and distant whaling and the rapid decline of sperm whales off the Galápagos Islands. Conservation Biology **11**:1387–1396.

Whitehead, H., and J. Gordon. 1986. Methods of obtaining data for assessing and modelling sperm whale populations which do not depend on catches. Reports of the International Whaling Commission (special issue) **8**:149–166.

Whitehead, H., and S. Waters. 1990. Social organisation and population structure of sperm whales off the Galápagos Islands, Ecuador (1985 and 1987). Reports of the International Whaling Commission (special issue) **12**:249–257.

SUPPLEMENTARY MATERIAL

Programs which are designed to make the first and third of the methods described in this paper easily applicable to a wide range of data sets are available in ESA's Electronic Data Archive: *Ecological Archives* E082-018.

APPENDIX

In this Appendix, I show that the parameter estimates obtained by maximizing the sum of several, not necessarily independent, log-likelihoods, each with the same parameters and expected parameter values, is approximately unbiased.

Suppose $\{L_i(p)\}$ are a set of, not necessarily independent, log-likelihood functions of the same parameter p. $L_i(p)$ is maximized at p_i, and the expected value of p_i is p_0 for all i. What is the expected value of p_A that maximizes $L_A = \sum L_i(p)$? Note that, because of nonindependence, L_A is not the likelihood function for the entire data set.

Using the Taylor-Maclaurin expansion,

$$L_A = \sum L_i(p_i) + (p - p_i)L_i'(p_i)$$
$$- (p - p_i)^2 L_i''(p_i)/2 + \cdots. \quad (A.1)$$

But $L_i(p_i)$ is fixed and $L_i'(p_i) = 0$, as p_i maximizes $L_i(p_i)$, so L_A is maximized by maximizing

$$L_A = \sum - (p - p_i)^2 L_i''(p_i)/2 + \cdots. \quad (A.2)$$

If the likelihood functions are approximately quadratic, then it is reasonable to ignore higher order terms. So L_A is then maximized by

$$p_A = \sum p_i L_i''(p_i) \Big/ \sum L_i''(p_i). \quad (A.3)$$

This is the intuitively reasonable mean of the individual estimates of p weighted by the inverse of their estimated variances (as $L_i''(p_i) \approx 1/\mathrm{var}(p_i)$). Then, as the expected value of each p_i is p_0, the expected value of p_A is approximately (as $L_i''(p_i) \approx L_i''(p_0)$):

$$p_A = \sum p_0 L_i''(p_0) \Big/ \sum L_i''(p_0) = p_0. \quad (A.4)$$

This argument can easily be extended to vector-valued parameters. Thus, if we sum several, not necessarily independent, log-likelihoods, each with the same expected parameter values, the parameter estimates obtained by maximizing this sum are approximately unbiased.