Trajectory-based approaches to estimating velocity and diffusion from tagging data

Clay E. Porch
Southeast Fisheries Science Center
National Marine Fisheries Service, NOAA
75 Virginia Beach Drive, Miami, Florida 33149

Abstract.—Several new models are developed to estimate the velocity and diffusion of a population from tagging data. The new estimators apply the inverse principle to the individual trajectories of recovered tags rather than to their local abundance. These models require fewer assumptions and less information than do published abundance-based methods. Techniques are presented for a variety of circumstances, and both discrete and continuous parameterizations of the velocity field are included. The sensitivity of the estimators to violations of the assumptions was examined numerically by using stochastic simulations. The results suggest that the estimators are fairly robust but may fail under certain conditions. Extensions to accommodate these situations are discussed.

Manuscript accepted 17 April 1995.
Fishery Bulletin 93:694–709 (1995).

Tagging experiments have often been used to delineate animal movements. Historically, many of the analyses were limited to graphical portrayals of apparent migration routes and simple measures of the net swimming speed. Beverton and Holt (1957) introduced more rigorous treatments of tagging data based on the pioneering work of Skellam (1951). They postulated that the Fickian diffusion equation would be an adequate model for the dispersion of fish due to random motions and developed a technique for estimating the diffusion coefficient from tagging data. Subsequently, Jones (1959, 1976) developed a simple estimation procedure that distinguished the diffusion and net drift of returned tags. Saila and Flowers (1969) proposed using a special case of the advection-diffusion equation (Fickian diffusion with constant velocity) to model fish migration and developed a numerical technique to estimate the diffusion coefficient and net drift. More recently, Sibert and Fournier (1994) advocated the use of a more general form of the advection-diffusion equation that allows for mortality and discrete changes in velocity among areas. They also developed a new estimation procedure based on fitting numerical predictions to the observed distribution of recovered tags. Similar methods have been applied to the movements of passive tracers in the ocean by Fiadero and Veronis (1984) and Wunsch (1989).

Beverton and Holt (1957) recognized that the solutions to advection-diffusion equations are greatly complicated by heterogeneous diffusion rates and irregular boundary conditions (e.g. coastlines). They suggested replacing the diffusion equation with a system of area-specific equations linked together by transfer coefficients that measure the movement across the boundary of adjacent areas (box models). This simple abstraction, as well as others like it, has received considerable attention in recent years, and a number of papers have dealt with estimating the transfer coefficients from tagging data (Beverton and Holt, 1957; Sibert, 1984; Hilborn, 1990; Deriso et al., 1991; Hampton, 1991; Schweigert and Schwarz, 1993; Kleiber and Fonteneau, 1994; Salvado, 1994). In principle box models are not very realistic because they assume that movement within and among boxes occurs instantaneously, but in practice they may approximate the dynamics well enough to be useful.

All of the aforementioned procedures (except Jones's) estimate the movement of a population from the local abundance of recovered tags. In contrast, the methods developed in this paper estimate movement from the trajectories of recovered tags. Strictly speaking, the new models address the advection (ex-

pected velocity) and diffusion of a population. As such, they should provide useful alternatives for estimating the movement parameters of the advection-diffusion equation or individual-based simulations. They are applicable to box models only to the extent that such box models are analogous to finite difference approximations to the advection-diffusion equation. (In that case the transfer coefficients can be written as functions of the local velocity and diffusion rates.)

This article is divided into five sections. The first section highlights the conceptual differences between trajectory-based and abundance-based estimators of velocity. The mathematical details of the proposed trajectory models are developed next. Section two focuses on the process of advection and section three on the process of diffusion. The performance of the models and their sensitivity to violations of the assumptions are evaluated by using stochastic simulations in the fourth section. Finally, some practical considerations and extensions to the methodology are discussed.

General concepts

Velocity, advection, and diffusion defined

The term 'velocity' refers to the speed and direction in which an object (tag) moves. Mathematically, the velocity U at position x and time t is defined by the differential equation

$$U(x,t) = \frac{dx}{dt}. \tag{1}$$

The map of U that assigns a velocity to each point in space and time is called the velocity field.

In theory, the position of any given tag at any given time can be predicted from its velocity history by integrating Equation 1. In application, however, it is usually more practical to describe the collective velocity histories of a group of tags in terms of a common expected component $u(x,t)$ and unique random components $u'(x,t)$. The expected component is known as the advection field, and the random component gives rise to the diffusion field. By using this description, the position x_{iT} of the i'th tag after T units of time can be expressed in the general form

$$x_{iT} = x_{i0} + \int_{t_{i0}}^{t_{i0}+T} u(x_t,t)dt + \int_{t_{i0}}^{t_{i0}+T} u'(x_t,t)dt, \tag{2}$$

where x_{i0} and t_{i0} are the initial position and time, respectively.

The first integral in Equation 2 determines the expected position of the tag, and the second integral determines the displacement of the tag relative to its expectation. In terms of a group of tags with common starting points, the first integral describes the advection of the group as a whole, and the second integral determines how the group spreads about its expected center of mass.

Approaches to estimating velocity

The purpose of estimating movement rates and other parameters from tagging data is usually to elucidate the behavior of a larger population—often within the context of managing that population. Any such application implicitly assumes that the tagged and untagged members of the population move the same way. This basic tenet is accepted throughout the remainder of this article. The discussion in this section focuses on the ancillary assumptions that different estimation approaches must make.

The classical formula for estimating the advection of a tag is

$$\hat{u} = \frac{1}{n}\sum_{i=1}^{n} \frac{x_{iT} - x_{i0}}{T_i}, \tag{3}$$

where n is the number of observations. Jones (1959, 1976) developed an alternative formula, which, in the present notation, is

$$\hat{u} = \frac{\sum_{i=1}^{n} x_{iT} - x_{i0}}{\sum_{i=1}^{n} T_i}. \tag{4}$$

The practicality of the estimators (Eqs. 3 and 4) is limited because for these the advection field is assumed to be relatively constant, which is unlikely over the temporal and spatial scales relevant to population management. Therefore, it will often be profitable to enlist a more dynamic model of advection. The parameters of the dynamic model can be estimated by minimizing an appropriate objective function of some measure of the effect of advection on the population. To date, the measure of choice for tagging data has been the local abundance of recovered tags. The measure used in this article is the trajectory of each individual tag.

The accuracy of any estimator depends on the assumptions behind the construction of the objective function. These involve assumptions regarding the probability density of the measure and the structure

of the measure's predictor. To illustrate, consider a least-squares objective function whose measure is the number N of tags recovered in each area a and season s:

$$\sum_s \sum_a (\hat{N}_{as} - N_{as})^2. \quad (5)$$

The predictor, $\hat{N}$, is a function of the parameters of the underlying population dynamics and advection models.

In theory, the parameter estimates that minimize Equation 5 are unbiased if the measure (N) is normally distributed with constant variance and the models behind the predictor are correct. It is generally recognized, however, that the probability density of recoveries is nonnormal and the variance in N may differ widely among areas and seasons, both of which imply that the least-squares solution is inappropriate. For this reason, most of the recent literature has favored maximum-likelihood solutions based on the multinomial distribution instead.

The condition of the predictor ($\hat{N}$) is more difficult to assess because it embodies a suite of assumptions regarding all relevant aspects of the population dynamics. The local abundance of tags may be affected by processes common to both the tagged and untagged populations (e.g. natural and fishing mortality) as well as by processes that are unique to the tagged population (e.g. tag-induced mortality, tag shedding, and failures to report recovered tags).

It would be advantageous to develop a predictor that does not need to account for all the complex processes affecting tag recoveries, but it is clear that a measure other than abundance must be used. The trajectories of individual tags, which can be predicted from their velocity history alone, is one such measure. Tag recovery rates are relevant to trajectories only in the sense that they determine those most likely to be represented in the sample. That is, recovery rates dictate the probability density of observed (recovered) trajectories, but not the predictor. This point, though subtle, has important implications with respect to relaxing the assumptions required to produce unbiased estimates of the advection field.

Consider that the expected position of a tag after liberty time T_i follows from Equation 2:

$$E[x_{iT}] = x_{i0} + \int_{t_{i0}}^{t_{i0}+T_i} u(x_t,t)dt. \quad (6)$$

The expected position of a recovered tag under the same conditions is

$$E_R[x_{iT}] = x_{i0} + \int_{t_{i0}}^{t_{i0}+T_i} u(x_t,t)dt + E_R\left[\int_{t_{i0}}^{t_{i0}+T_i} u'(x_t,t)dt\right],$$

where the subscript R indicates that the expectation includes recovered tags only. The second integral dropped out of the unconditional expectation in Equation 6 because u' is, by definition, a random variable with mean equal to zero. The same would not generally be true of the expectation of recovered tags because some vectors of u' may be more likely to be recovered than others—changing the probability density in some unknown way.

Suppose there exists an objective function $O[x,\hat{x}]$ (maximum likelihood or otherwise) that can produce unbiased estimates of $E[x_{iT}]$ from a random sample of all potential trajectories x_{iT}. (The construction of this function will be discussed later.) The same objective function will also produce unbiased estimates from a random sample of recovered tags provided

$$E_R\left[\int_{t_{i0}}^{t_{i0}+T_i} u'(x_t,t)dt\right] = 0. \quad (7)$$

This constraint is satisfied if either u' is everywhere identically zero or the probability of recovering a tag is independent of its velocity. The latter condition is effectively equivalent to assuming that the processes that influence tag recovery are homogeneous in space and time. It satisfies Equation 7 because it implies that the relative likelihood of observing any given displacement depends solely on the probability density of u'. By definition, the expectation of u' at every point is zero and therefore the expectation of the integral sum of u' is also zero.

It has been shown that, subject to Equation 7, trajectory-based estimators can provide unbiased estimates of the population advection field without recovery rates having to be considered. One can imagine many practical situations where Equation 7 would be approximately satisfied. The spatial and temporal distribution of recovery rates would not normally be a significant factor in experiments involving radio-tracked drifter buoys or ultrasonic tags. Similarly, variations in the velocity of individuals might be expected to be small compared with the average velocity of a population undergoing a seasonal spawning migration. Where Equation 7 is not met, however, tags moving at different velocities may not be equally represented in the recovered sample.

If, for example, faster tags were more likely than slower tags to move into a region where recovery rates are high, then the speed of the advection field would be overestimated.

At this point it seems convenient to examine the proposed trajectory models in closer detail, noting that recovery factors other than advection do not need to be considered when Equation 7 is satisfied. The more complicated matter of accounting for variations in recovery rates when Equation 7 is not met is deferred to the Discussion section at the end of this article.

Trajectory-based predictors for estimating advection

This section focuses on developing the predictor for the directed (advective) component of motion from the trajectories of recovered tags. The remaining portion of the objective function, which quantifies the differences between the observed and predicted values of the measure, depends on the nature of the probability density and is discussed in detail in the subsequent section entitled "Diffusion and the objective function."

The general form of the trajectory predictor may be written

$$\int_{x_{i0}}^{\hat{x}_{iT}} dx = \int_{t_{i0}}^{t_{i0}+T_i} \hat{u}(x_t,t)\,dt, \tag{8}$$

where $\hat{x}_{iT}$ = predicted position of tag i after time at liberty T;
$\hat{u}$ = estimated advection field;
x_{i0} = initial position of tag i; and
t_{i0} = initial date of tag i.

In order to use this prediction equation, one must be able to evaluate the integral on the right. There are two ways to address this problem. One way is to break the temporal and spatial domain into small strata where the advection rates are approximately constant and then to assemble a picture of the large-scale advection field in piece-wise fashion. The other way is to define explicitly a dynamic model of the advection field and to evaluate the integral directly. Each approach is developed in a separate subsection below.

Piece-wise models

This approach seeks to assemble a picture of the overall advection field from estimates of the advection fields in smaller strata. An independent estimate of the average advection in each space and time strata can be obtained from any tag that has remained in that strata the entire time between its release and recovery (or between any two position updates) by using the formula

$$u_i = \frac{x_{iT} - x_{i0}}{T_i}.$$

If observations are available for most of the strata of interest, the entire advection field can be parameterized quite nicely by using a two-way analysis of variance (ANOVA) model:

$$u_i = \bar{u} + A_a + S_s + I_{as} + \varepsilon_i, \tag{9}$$

where u_i = the observed velocity of tag i;
$\bar{u}$ = the overall mean velocity;
A_a = the main effect of area a on u;
S_s = the main effect of season s on u;
I_{as} = the area/season interaction effect; and
ε_i = the error associated with tag i.

In two spatial dimensions a separate ANOVA would apply to each velocity component:

$$u_i = \bar{u} + A_{au} + S_{su} + I_{asu} + \varepsilon_{iu},$$
$$v_i = \bar{v} + A_{av} + S_{sv} + I_{asv} + \varepsilon_{iv},$$

where u_i = the observed velocity in the direction of the first dimension;
v_i = the observed velocity in the direction of the second dimension;
$A_{a\_}$ = the main effect of area a on u or v;
$S_{s\_}$ = the main effect of season s on u or v;
$I_{as\_}$ = area/season interaction effect on u or v; and
$\varepsilon_{i\_}$ = the errors associated with tag i.

Interpolation routines other than ANOVA may also be used to describe the overall advection field, but ANOVA provides a convenient framework for testing whether the advection rates vary among strata. Such tests are valid if the velocity variances are the same in all strata; otherwise one must employ an equivalent nonparametric approach.

ANOVA or other interpolation algorithms are well suited to situations where the positions of tags can be updated frequently. They are especially promising for programs that employ remote tracking devices such as radio or ultrasonic tags (e.g. Quinn, 1988; Hines and Wolcott, 1990; Schulz and Berg, 1992). In some cases the biology of the organism may even permit effective visual tracking (e.g. Stachowitsch, 1979). The ANOVA approach is less suited to

conventional tagging programs because the ability to record the position of the tag is largely beyond the control of the investigator. In this situation a sufficient number of tags would need to be released in each season and area strata to ensure that at least a few would be recovered before straying into other strata. The recovery times must also be short enough to avoid biasing the results toward slower moving tags.

A more flexible estimation procedure, which is capable of incorporating trajectories that reflect the combined effects of several different movement patterns, can be derived by reformulating Equation 8 as

$$\int_{x_{i0}}^{x_{iT}} \frac{1}{u(x,t)}dx = \int_{t_{i0}}^{t_{i0}+T_i} dt. \qquad (10)$$

If $u(x,t)$ is held constant within specific space and time strata but allowed to vary among strata, then Equation 10 may be piece-wise integrated. The solution simplifies to

$$x_{s+1} = \left(t_{s+1} - t_s - \frac{B_{\alpha+1} - x_s}{u_{\alpha s}} - \frac{B_{\alpha+2} - B_{\alpha+1}}{u_{\alpha+1,s}} - \cdots\right)u_{\Omega,s} + B_{\Omega}, \qquad (11)$$

when $u > 0$, and

$$x_{s+1} = \left(t_{s+1} - t_s - \frac{B_{\alpha} - x_s}{u_{\alpha s}} - \frac{B_{\alpha-1} - B_{\alpha}}{u_{\alpha-1,s}} - \cdots\right)u_{\Omega,s} + B_{\Omega+1}, \qquad (12)$$

when $u < 0$.

Here u_{as} = velocity in area a during season s;
B_a = boundaries of the areas (see Fig. 1);
α = first area occupied by the tag during season s;
Ω = last area occupied by the tag during season s;
t_s = date at the onset of the s'th season; and
x_s = position of tag at onset of s'th season.

With these recursions, the position of the tag at the end of each season can be computed from the tag's position at the end of the previous season by using an estimate of the advection field. This procedure can then be applied sequentially to compute the expected position of the tag at the date when it was recovered from its initial position. The "best" estimates of the strata-specific advection rates would minimize some objective function of the differences between the observed recovery positions and those predicted with the recursions. (The appropriate objective function will be discussed in the section on diffusion.)

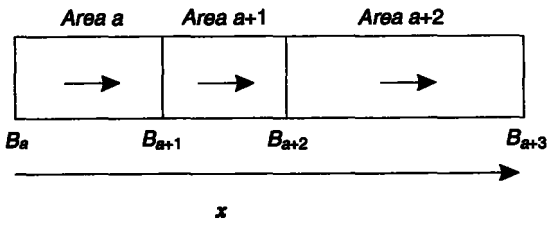

Figure 1
Schematic of a region divided into three areas illustrating the definition of the area boundaries, B.

The sequential procedure itself is accomplished by first determining the starting season s_0 and starting area a for each tag from the date and position where it was released. The position of the tag at the end of the first season (start of season s_0+1) can then be obtained from the recursions, by replacing x_s and t_s with the position and date of release. The recursions are then applied as given until the last season, which is determined by the recovery date. Finally, the formula for estimating the recovery position is obtained by replacing t_{s+1} in the recursions with the recovery date.

Continuous models

The approach proposed in this section involves developing an adequate continuous model of the advection field, $u(x,t)$, and solving Equation 8 for x as a function of t. To illustrate, consider a fish population migrating out of a basin. Suppose each fish swims initially at speed u_0 in the positive x direction and increases its speed as it proceeds. Further, suppose that periodically the fish are either helped or hindered by sinusoidal oscillations in the water currents. A reasonable model for the velocity of the fish might be $u(x,t) = u_0 + ax + b\sin[ct]$. The solution is

$$x = -\frac{u_0}{a} - b\frac{a\sin[ct] + c\cos[ct]}{a^2 + c^2} + \gamma e^{at}, \qquad (13)$$

where

$$\gamma = \left(x_0 + \frac{u_0}{a} + b\frac{a\sin[ct_0] + c\cos[ct_0]}{a^2 + c^2}\right)e^{at_0}.$$

The best estimates of the parameters u_0, a, b, and c would be those that minimize some appropriate func-

tion of the difference between the observed and predicted positions of all the tags.

In practice, models capable of effectively capturing the dynamics of the advection field may not be simple enough to admit an analytical solution. The obvious alternative is to evaluate the right-hand integral in Equation 8 numerically. Press et al. (1986) contributed an excellent review of many of the more powerful numerical integration algorithms.

It may also happen that the system is not understood well enough to construct a detailed model of tag motion. In that case there is no equation to integrate, analytically or otherwise. One option is to resort to the piece-wise methods described earlier. There is, however, a second option available provided the effects of space and time on advection are separable, i.e. provided Equation 1 may be recast as

$$\frac{dx}{g[x]} = h[t]dt, \tag{14}$$

where g and h are functions of space and time, respectively. In this case the right and left sides of Equation 14 may be integrated separately:

$$\int_{x_0}^{x_T} \frac{1}{g[x]} dx = G[x_T] - G[x_0] \tag{15}$$

and

$$\int_{t_0}^{t_0+T} h[t]dt = H[t_0 + T] - H[t_0]. \tag{16}$$

It may not always be possible to derive the functions G and H from g and h, but g and h can always be derived from G and H (provided they are continuous on the interval). To the extent that Equations 15 and 16 are equivalent, the suggestion is to pose flexible functions for G or H and to fit them to the data by minimizing their differences. The least-squares fit, for example, would minimize the quantity

$$\sum_i \left(G[x_{iT}] - G[x_{i0}] - H[t_{i0} + T_i] + H[t_{i0}]\right)^2.$$

The corresponding advection field would be obtained by the formula

$$\hat{u}[x,t] = \frac{d\hat{H}}{dt}\left(\frac{d\hat{G}}{dx}\right)^{-1}$$

(a proof is given in the Appendix).

The efficacy of the separability procedure will depend on the behavior of the functions chosen to represent G and H. Although the functions must be flexible enough to accommodate a wide range of movement possibilities, they must also be compatible with the search routine used to find the best estimates of the parameters. When polynomials are used, for example, the search tends to converge on the trivial solution $G = H = 0$. Depending on the search algorithm, it may be possible to impose constraints that circumvent such problems. In any case, it may help to use a theoretical model for either the spatial or temporal aspect of the problem. For example, if there is evidence to suggest that the pattern of motion is periodic, one might write

$$\int_{x_0}^{x_T} \frac{1}{g[x]} dx = \int_{t_0}^{t_0+T} \sin[c(t-d)]dt,$$

where the frequency (c) and offset parameter (d) are known. The corresponding least-squares function to be minimized is

$$\sum_i \begin{pmatrix} G[x_{iT}] - G[x_{i0}] - \cos[c(t_{i0} + T_i - d)] + \\ \cos[c(t_{i0} - d)] \end{pmatrix}^2.$$

The function G would be some arbitrary, but flexible, function.

Diffusion and the objective function

This section examines the random (diffusive) component of tag motion. There are two perspectives from which to do this: the absolute sense (the displacement of individuals from their expected positions) and the relative sense (the displacement of individuals in a patch relative to one another). The distinction is important because relative diffusion includes only those physical processes acting within the patch and implicitly excludes processes that might cause the patch itself to vary from its expected path. Absolute diffusion, on the other hand, includes random processes operating on all scales. The trajectory approaches espoused in this article model each tag without regard to its proximity to other tags; therefore the statistics they produce are not relevant to the diffusion of a patch. Accordingly, the remainder of this discussion will focus on diffusion in the absolute sense.

Measures of diffusion

Two common measures of absolute diffusion are mean-square dispersion and absolute diffusivity.

Mean-square dispersion is defined as the average of the squared deviations of the tags from their expected positions. It is usually expressed as a function of elapsed time T:

$$\chi^2_{iT} = \frac{1}{n}\sum_{i=1}^{n}(x_{iT} - \hat{x}_{iT})^2, \quad (17)$$

where $\hat{x}_{iT}$ is the predicted position given the estimated advection field, and n is the number of tags in the sample. The absolute diffusivity is defined as one half the time rate of change of the mean-square dispersion:

$$K_T = \frac{1}{2}\frac{d(\chi^2_T)}{dt}.$$

Taylor (1921) showed that the mean-square dispersion of particles in a homogeneous random field increases linearly with time, provided that the particles have been at large long enough for their present motions to become statistically decorrelated from their initial motions:

$$\chi^2 = \beta T.$$

By definition then, the diffusivity associated with homogeneous random motions is constant at $\beta/2$, and the dispersion of particles is governed by the advection-diffusion equation. Under these conditions the diffusivity is synonymous with the "diffusion coefficient" of the Fickian advection-diffusion equation.

Implications for the objective function

The time-dependent nature of the mean-square dispersion has important implications for the nature of the objective function used in estimating the parameters of the advection field. When the diffusivity is constant, the positions of tags with initial conditions i will follow a normal distribution with mean $E[x_{iT}]$ and variance βT (Okubo, 1980). In addition, if the recovery positions are reported with normally distributed errors, the observed positions of the tags will follow a normal distribution with variance $\beta T+\sigma^2$. The maximum-likelihood estimates are therefore those that minimize the weighted least-squares formula

$$\sum_i \frac{(x_{iT} - \hat{x}_{iT})^2}{\beta T_i + \sigma^2}, \quad (18)$$

where $\hat{x}_i$ is the predictor defined by Equation 8. In effect, Equation 18 prevents tags that are expected to have large random displacements from dominating the analysis by down-weighting them according to their time at large.

The variance parameters in Equation 18, β and σ^2, must either be known or estimated as part of the search. However, if there is some evidence that the observation errors are much larger than the displacements attributable to random motions, then Equation 18 reduces to ordinary least squares. Similarly, if the random displacements are much greater than the observation errors, then Equation 18 reduces to

$$\sum_i \frac{(x_{iT} - \hat{x}_{iT})^2}{T_i}. \quad (19)$$

There are many practical circumstances where it may be reasonable to assume a constant diffusivity and to apply Equation 18. Observations in the open ocean suggest that the turbulent motions in many regions are approximately homogeneous (e.g. de Verdiere, 1983; Krauss and Boning, 1987; Figueroa and Olson, 1989; Poulain and Niiler, 1989). Moreover, Porch (1993) points out that random walk models of fish movement also exhibit mean-square dispersions that increase in proportion to time (provided swimming speed does not increase substantially during the time period of interest).

The maximum-likelihood formulation for an inhomogeneous diffusion field is unclear. Equation 18 may be acceptable where the random displacements of the tags increase monotonically with time, but it will not be acceptable in all cases. When animals are aggregating to spawn or feed, for example, the effective diffusivity would be zero or negative. Likewise, the presence of coastal boundaries complicates the matter because the diffusivity in one direction is zero. For this reason, it may be more prudent to consider methods that are robust to inhomogeneous variances, such as least-median-of-squares regression (Rousseeuw, 1984) and least-absolute-value regression (Bloomfield and Steiger, 1983).

Estimating diffusivity

As mentioned previously, the position of a tag at liberty for time T in a homogeneous diffusion field follows the normal distribution with mean $E[x_{iT}]$ and variance βT. This implies that the displacement of the tag relative to its expected position (D) is also normally distributed with mean 0 and variance βT. The squared displacements are therefore gamma-distributed with parameters 1/2 and $1/(2(\beta T+\sigma^2))$. Accordingly, the maximum-likelihood estimates for β and σ^2 are those that satisfy the constraints

$$\sum_i \frac{T_i(D_i^2 - \beta T_i - \hat{\sigma}^2)}{(\beta T_i + \hat{\sigma}^2)^2} = 0$$

and

$$\sum_i \frac{D_i^2 - \hat{\beta} T_i - \hat{\sigma}^2}{(\hat{\beta} T_i + \hat{\sigma}^2)^2} = 0,$$

where $D_i = x_{iT} - \hat{x}_{iT}$. These equations must be solved numerically. When σ^2 is negligible, however, the maximum likelihood estimator for β reduces to

$$\hat{\beta} = \frac{1}{n} \sum_{i=1}^{n} \frac{D_i^2}{T_i}.$$

Less efficient estimates of β and σ^2 can be obtained from a linear regression of the squared displacements on the time at liberty:

$$D_i^2 = mT_i + b + \varepsilon_i,$$

where ε is a random deviation term. The slope of the regression m is an estimate of β and the intercept b is an estimate of σ^2.

Estimator performance

This section presents the results of a series of stochastic simulations designed to test the efficacy of trajectory-based estimators. Although no attempt was made to exhaust the possibilities, a sufficiently broad range of conditions was examined to validate the methodology. The experimental design and results are discussed in the subsections below.

Experimental design

A factorial design was used to study the behavior of the estimator in response to the advection field, level of diffusivity, distribution of recovery rates, and number of tags recovered. Two hundred test data sets were generated for each possible combination of factors. The estimators were then applied to each of the test data sets.

Response variables The accuracy of the estimators was quantified by the percent error of the parameter estimates ($\hat{\theta}$) relative to the actual values (θ_{true}):

$$PE = \frac{1}{200} \sum_{j=1}^{200} \frac{\hat{\theta}_j - \theta_{true}}{\theta_{true}} 100\%.$$

The precision of the estimators was quantified by the coefficient of error (CE),

$$CE = 100\% \sqrt{\frac{1}{200} \sum_{j=1}^{200} \left\{ \frac{\hat{\theta}_j - \theta_{true}}{\theta_{true}} \right\}^2}.$$

The coefficient of error is analogous to the familiar coefficient of variation except that the true values of the parameters are used in place of the average of their estimates. A low CE, therefore, implies that the estimates are both unbiased and precise.

Factor levels Two advection models were considered: the sinusoidal model $dx/dt = u_0 + ax + b\sin[ct]$, discussed earlier in connection with Equation 13, and a discrete model with two areas and semi-annual seasons. The parameters of the first model—u_0, a, b, and c—were valued at 8.6 km·day^{-1}, 0.004 day^{-1}, 17.3 km·day^{-1}, and $2\pi/365$ day^{-1}, respectively. The parameters of the second model are the area and season-specific constant advection rates. The rates in areas 1 and 2 were set equal to 5 and 10 km·day^{-1} during the first season and –7 and –4 km·day^{-1} during the second season. Area 1 extended from negative infinity to 1,000 km and area 2 extended from 1,000 km to positive infinity.

Two diffusivity ($\beta/2$) levels, 0.95 and 822 km^2·day^{-1}, were examined. These levels were derived by assuming that tagged fish move according to the bilateral random walk model (see Porch, 1993) with an average speed of 0.5 or 6 meters per second and change direction an average of once per minute or once every five minutes, respectively.

The effects of variations in tag recovery rates were evaluated by dividing the relevant spatial domain into two zones and by varying the likelihood of recovering a tag between them. Three such scenarios were considered. In the first, the probability of recovery was the same in both zones. In the second, the probability of recovery was ten times higher in zone A than in zone B (1.0 versus 0.1). In the third scenario, no tags were recovered in zone B. The boundary separating recovery zones A and B differed with the advection models. The demarcation point was x = 400 km in the sinusoidal model and x = 0 km in the discrete model.

Test data Each test data set was generated by simulating the individual paths of a prescribed number of tags (n). The release positions were randomly assigned values between 0 and 200 km when the sinusoidal advection model was used and were between –1,000 and 1,000 km when the discrete model was

employed. The release dates were randomly assigned values between day 0 and day 365. These choices mimic an opportunistic tagging program where the number of tags released is relatively constant during the year. The recovery dates were obtained by randomly selecting the liberty times from an exponential distribution with parameter Z (instantaneous mortality rate) equal to 0.4 yr^{-1}.

The displacement of each tag due to advection was computed from its release position and from release and recovery dates by using the solutions to the determinate advection models described above. The solution to the sinusoidal model is given by Equation 13, and the solution to the discrete model is given by Equations 11 and 12. The diffusive effect of random motions was then simulated by adding a random normal deviate with mean 0 and variance βT km^2. Next, an acceptance-rejection criterion was invoked to determine whether or not the tag would be recovered. Candidate tags located in recovery zone A were unconditionally accepted, but candidate tags located in zone B were accepted only with probability P (=1.0, 0.1, or 0). This was done by generating a uniform random number between 0 and 1 and by excluding the tag if that number was greater than the prescribed recovery probability.

The process described above was repeated until a total of n recovery positions were accepted. Normally distributed errors (with variance 25 km^2) were then added to each of the accepted release and recovery positions to simulate imprecise position reporting. In this way an artificial sample of n tag recoveries was created.

Estimation The predictors were fitted to the test data by minimizing the weighted least-squares surface described by Equation 19. The predicted positions were calculated by substituting estimates of the parameters into the same advection equations used to generate the data. This allowed the analyses to focus on the interactions of recovery rates and velocity variance without the confounding effects of model misspecification.

The minimization was accomplished by using the Nelder-Mead simplex algorithm AMOEBA (Press et al., 1986), which, although slower than derivative-based methods such as Marquardt's algorithm, is less sensitive to the discontinuities in the solution surface associated with discrete advection models. Heavy penalties were imposed to prevent the search from extending beyond the bounds of a reasonable domain. For example, the maximum possible sustained speed of a migrating tuna might be the sum of the cruising speed of the fish and the maximum speed of the water currents.

The AMOEBA search was restarted at the point P_0, where a minimum had been found, to avoid local anomalies in the solution surface. Subsequent "restarts" continued until five consecutive sets of parameter estimates differed by less than one percent. New vertices were selected for each restart by using the formula

$$P_{ij} = P_{0j} e^{0.5\lambda\delta_i} \; (i, j = 1, \ldots, \omega),$$

where P_{ij} is the value of the j'th coordinate (parameter) in the i'th vertex of the initial simplex, λ is a standard normal variate, and δ_i is equal to one if i equals j and zero otherwise.

Results

This section is divided into two parts, each focusing on the results pertaining to one of the two types of advection models.

Sinusoidal model

The estimation procedure generally behaved very well when the frequency (c) of the sinusoidal oscillations was known and the diffusivity was low (0.95 $km^2 \cdot day^{-1}$). The CE's, which reflect both accuracy and precision, were very low regardless of the distribution of recovery rates (Fig. 2). For the most part, the estimator continued to perform well at high diffusivities (822 $km^2 \cdot day^{-1}$). When the recovery probabilities were the same in both zones, the estimates were unbiased and the CE's rapidly decreased with increasing sample size to less than 10 percent. The estimates were only slightly biased and similarly precise even with a tenfold difference between the recovery probabilities. It was only when no tags were recovered in zone B (i.e. beyond the 400-km demarcation) that the estimates were significantly biased.

The trends were very similar to those described above when the frequency parameter c was estimated along with the other three parameters. A few of the runs, however, failed to converge to acceptable solutions—the weighted least-squares function being an order of magnitude greater than that expected, given the known diffusivity. This problem is not surprising considering the oscillatory nature of any periodic surface. Even if the true values of the other parameters were known, the surface map of the objective function would be characterized by local peaks and valleys that vary with the estimate of c. This behavior is demonstrated by a simplified model of the residual sum of squares (Fig. 3). Although the amplitudes of the peaks and valleys in the more complicated

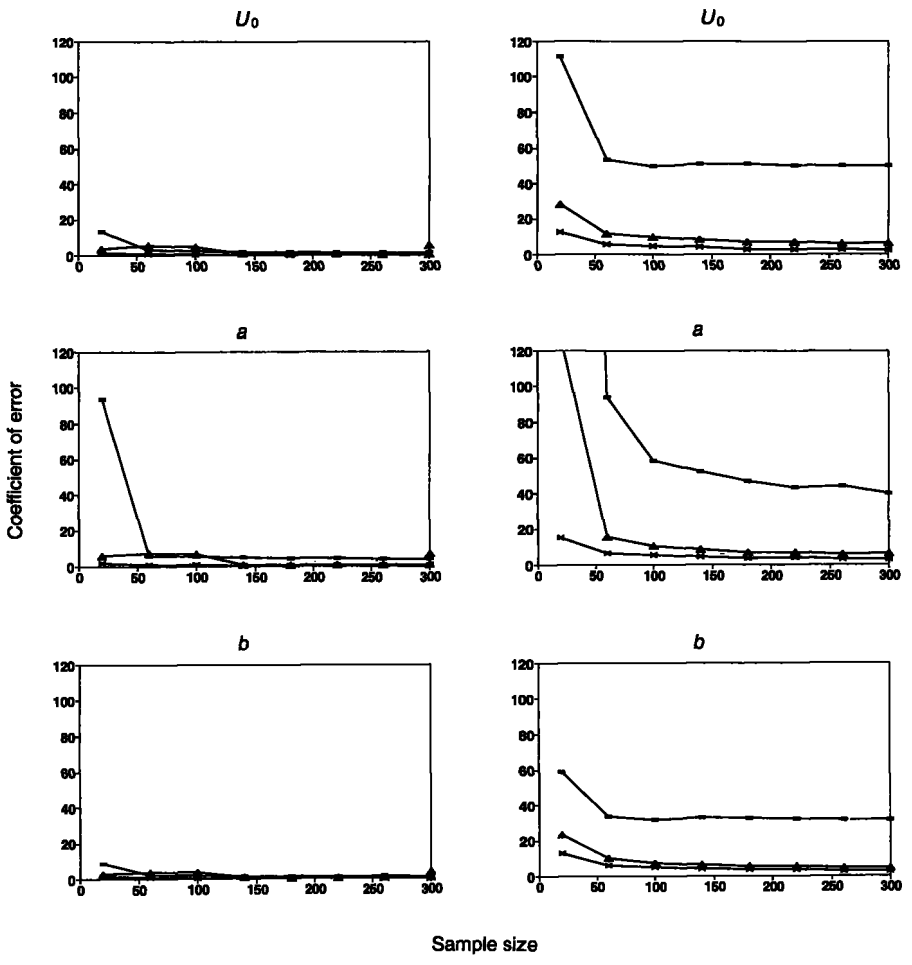

Figure 2

Coefficients of error of the estimates for the parameters of the sinusoidal model. The graphs on the left give the coefficients of error (CE's) under weak diffusivity and the graphs on the right give the CE's under strong diffusivity. The three curves in each graph correspond to zone B recovery probabilities of 0.0 (squares), 0.1 (triangles), and 1.0 (crosses).

model vary (the lowest valley presumably occurring in the vicinity of the true values of the parameters), several of the valleys may be deep enough such that the estimation routine finds a false global minimum.

The convergence problem was eliminated when the search was confined to a relatively restricted range of periodicities and was supplied with good initial guesses. In practice, adequate initial guesses and relatively narrow ranges can usually be deduced even from anecdotal data, therefore this should not prove too serious a limitation to the method. In cases where the periodicity is totally unknown, one should search the entire feasible domain with as fine a resolution as possible.

Piece-wise discrete model

The coefficients of error associated with each parameter were, for the most part, very low when the diffusivity was low (Fig. 4). However, the estimates pertaining to area 2 were highly biased and imprecise when the probability of recovering a tag was 0.0 in zone B. This is not unexpected given that excluding recoveries in zone B (x>0 km) all but eliminates the possibility that any of the recovered tags would ever have encountered area 2 (x>1,000 km). Thus, there is essentially no information on the advection in area 2 and the estimation routine fails. A similar result was not observed when the recovery probabil-

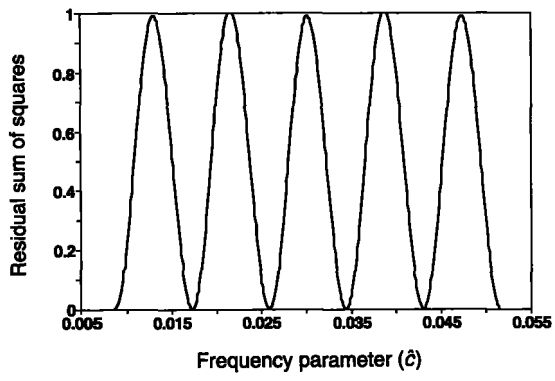

Figure 3

Map of the function $(\sin[ct] - \sin[\hat{c}t])^2$, where c is the true value of 0.0172 day^{-1} and $\hat{c}$ represents an estimate of c.

ity in zone B was 0.1 because a substantial fraction of the tags in each sample were still recovered in advection area 2 (see Fig. 5). Some of the tags recovered in area 1 undoubtedly passed through area 2 at some point as well, further augmenting the amount of information pertinent to estimating the advection in area 2.

The estimation procedure did not perform nearly as well at high diffusivities as it did at low diffusivities except when the recovery rates were the same in both areas (1.0). In that case the estimates remained unbiased and relatively precise—the CE's having dropped rapidly with sample size to less than twenty percent (Fig. 6). The CE's increased dramatically when the recovery rates differed between zones, mostly reflecting the corresponding increase in bias (Fig. 7).

The trends in the CE's also indicate that localized increases in recovery rates may improve the precision of local estimates at the expense of the precision of estimates for the other areas. The CE's in area 2 went down with increased recovery rates in zone B, but the CE's in area 1 went up. This effect, however, is largely an artifact of keeping the sample size fixed; increasing the fraction of recoveries near area 2 directly decreases the effective sample size for area 1 and vice-versa. In reality, local increases in recovery rates should add to the number of local recoveries more than they subtract from the number of recoveries elsewhere, unless the overall recovery rates are very high and there is a great deal of mixing between zones.

Discussion

The methods advanced in this article are fundamentally different from those cited previously because they predict trajectories rather than local abundance. Abundance-based and trajectory-based approaches both assume that tagged and untagged populations move the same way, but they differ in the ancillary assumptions they make. Abundance-based estimators, by their very nature, must enumerate a very large number of assumptions regarding processes that could affect the local abundance of the tags. In contrast, the trajectory-based estimators discussed so far disregard everything but velocity.

It was demonstrated earlier that trajectory-based estimators can produce unbiased estimates of the advection field for the population in general provided that either the tag recovery rates are homogeneous in space and time or the velocity variance is small. If neither of the above conditions are met, the estimates may be biased because faster individuals are more likely to move through regions with different recovery rates and may, therefore, be overrepresented or underrepresented in the sample. This is true even if the data are obtained from archival tags. The simulation experiments, however, indicated that the bias may not be too severe unless the recovery rates differ a great deal between areas. Moreover, recovery rates generally will not be a factor when the tags are tracked by radio, ultrasonic, or visual means.

These conclusions suggest that trajectory-based methods are appropriate for many real data sets. In addition, in many cases they may be much more practical than abundance-based methods. Although all of the derivations so far have been in one dimension, it is relatively simple to extend the methodology to include two or three dimensional movements, as is done in the second subsection. Finally, a third subsection is devoted to discussing the possibilities for adjusting the estimators to account for strongly inhomogeneous recovery rates.

Practical utility of trajectory-based estimators

Trajectory-based estimators have several advantages over their abundance-based counterparts. First, whereas trajectory models operate in continuous space and time, tag abundance models are obliged to operate in discrete space and time (one cannot speak meaningfully of the numbers of tags recovered at infinitesimal points). Thus, abundance models suffer from truncation errors that occur when different positions are lumped in the same category. If the time and space grid is fine enough, the truncation error will be minimal, but there may then arise the practical estimation problem of having very few observations in any given space-time category.

A second practical advantage relates to the choice of models. Both trajectory-based and abundance-

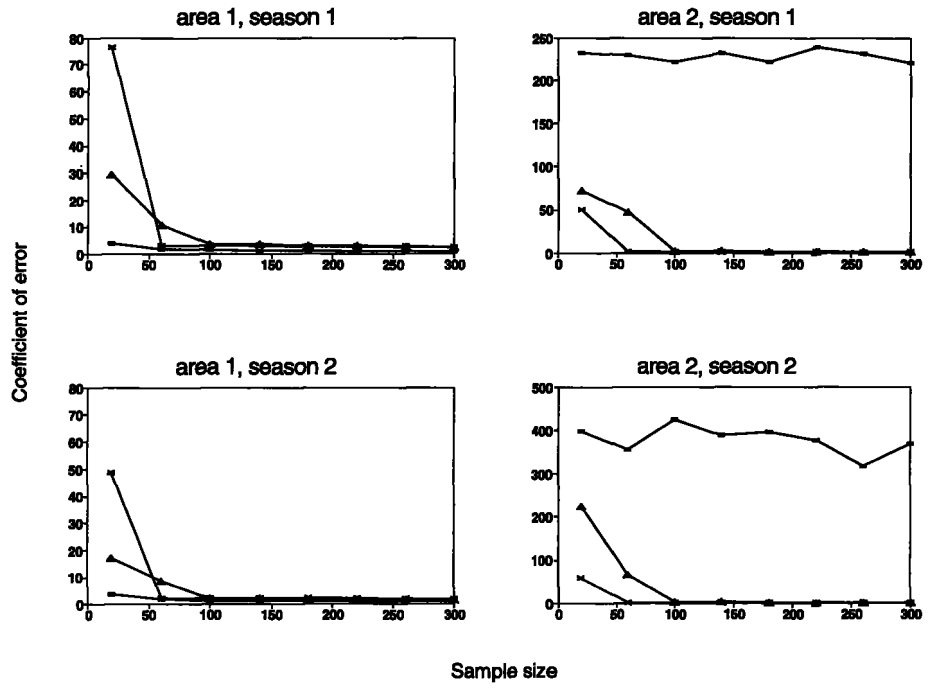

Figure 4

Coefficients of error of the estimates for the parameters of the discrete model under weak diffusivity. The three curves in each graph correspond to zone B recovery probabilities of 0.0 (squares), 0.1 (triangles), and 1.0 (crosses).

based methods must prescribe a velocity model, but the latter must also specify models for a great many other processes. Even if the velocity model is correct, abundance-based estimates may still be biased if any of the models for the other components are misspecified. Furthermore, because abundance-based methods estimate many parameters, they tend to require a large amount of data. In the absence of such a large data base, trajectory-based approaches may be the only reasonable option.

Another attractive feature of trajectory-based approaches is that they naturally accommodate data from archival tags. The chronological sequence of n position updates from each tag can be treated as though it were a sample of n independent tags with short liberty times. Conventional abundance-based methods cannot take advantage of this additional information.

Trajectory-based approaches do have their limitations. One is that they are not useful for assessing aspects of the population dynamics other than velocity. This point is especially important when the behavior of the population is being investigated within a management context. Under these circumstances abundance-based methods would seem to be the more viable option because they can, at least in principle, be formulated to estimate any relevant parameter. The enormous number of parameters required of such models, however, compromises the efficiency of the parameter search. Moreover, abundance-

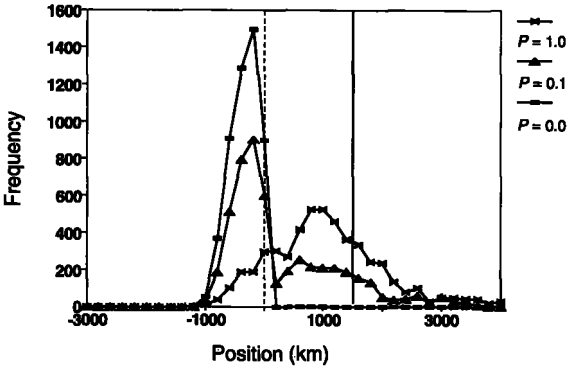

Figure 5

Distribution of 2,000 recovery positions generated by using the discrete advection model and weak diffusivity. The legend refers to the probability (P) of recovering a tag in zone B (right side of dashed vertical line). The solid vertical line marks the boundary between areas 1 and 2.

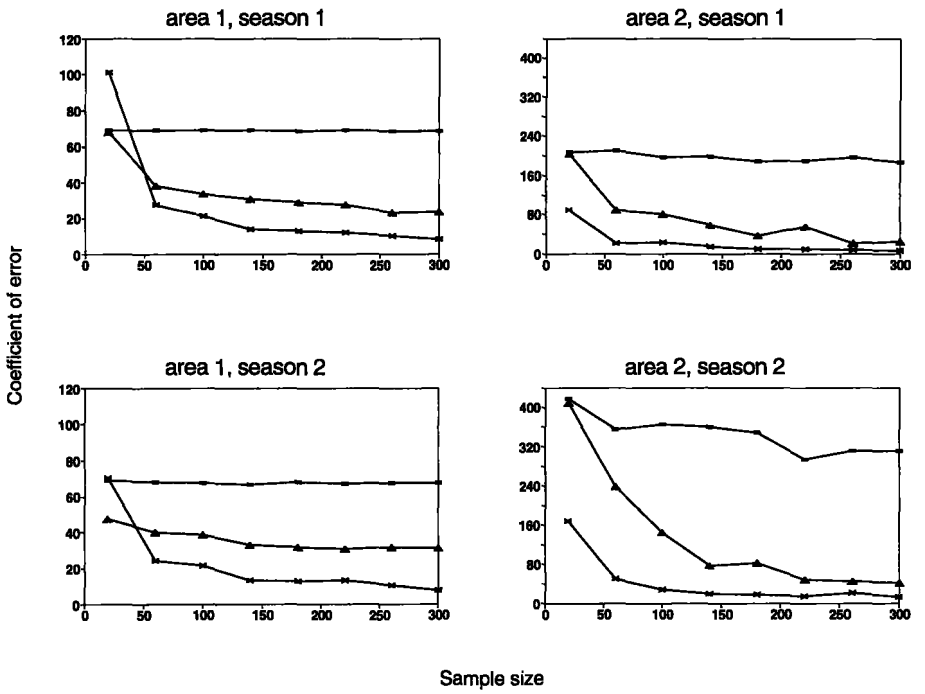

Figure 6

Coefficients of error of the estimates for the parameters of the discrete model under strong diffusivity. The three curves in each graph correspond to zone B recovery probabilities of 0.0 (squares), 0.1 (triangles), and 1.0 (crosses).

based estimates of movement and of certain other parameters, such as fishing mortality, are highly correlated (Hampton, 1991; Porch et al. in press; Aldenberg[1]). A number of competing hypotheses may return similarly low values of the objective function. Trajectory-based methods may prove very useful in this regard by supplying independent estimates of the velocity field that can be fixed in the abundance-based model. This would both reduce the dimensions of the search and eliminate some of the correlation problems. Another possibility might be to combine the abundance and trajectory-based approaches by including both formulations in the objective function.

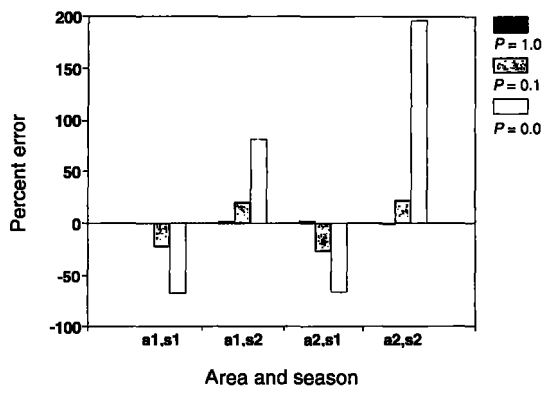

Figure 7

Percent error of the estimates for the parameters of the discrete model under strong diffusivity. The legend refers to the probability (P) of recovering a tag in zone B.

The most important limitation of trajectory-based estimators is that there is no guarantee that they are unbiased unless either the diffusive displacements are negligible relative to the advective displacements or the recovery rates are fairly homogeneous. Fortunately, the validity of the first assumption can be inferred rather easily from the output of the model itself. Estimates of the mean-square dispersion, which reflects the level of velocity variance, can be obtained by using Equation 17. If the square-root of the estimated mean-square dispersion is negligible compared with the actual displacement of the

[1] Aldenberg, T. 1975. Virtual population analysis and migration: a theoretical treatment. ICES Working Document (Counc. mtng.) 1975/F, 32 p.

tag, then the estimated advection field should be relatively accurate. Otherwise the estimates may be biased, either because the velocity variance is large or because the advection model is poorly specified.

The simulation results revealed that the estimators were essentially unbiased when the low level of diffusivity was used in generating the data. The root-mean-square displacements associated with this level of diffusivity were on the order of two percent of the advective displacements. This suggests that random displacements of at least two percent, and perhaps much larger, can be considered "negligible." Simulations that are tailored to specific situations are recommended to determine more accurately the tolerance of any particular application.

Extensions to multiple dimensions with boundaries

The mathematical derivations in this paper were developed in one dimension and without regard to barriers, primarily to simplify the presentation. It may sometimes be convenient to describe tag motion in this manner, such as when the tags are embedded in a major ocean current. Otherwise, the methods can easily be extended to accommodate more complicated scenarios.

Multiple dimensions can be incorporated by constructing dimension specific components in the objective function, e.g.

$$\sum_i \frac{(x_i - \hat{x}_i)^2}{\beta_x T_i} + \frac{(y_i - \hat{y}_i)^2}{\beta_y T_i},$$

where β_x and β_y are the diffusivity parameters in the two-dimensional space spanned by the coordinates x and y. The predicted positions are then obtained by integrating the differential equations describing the advection field in the two-dimensional space. Although multidimensional equations of motion are typically more difficult to solve analytically than their one-dimensional counterparts, they can always be integrated numerically. The other alternative is to use discrete approximations, in which case the motion along each dimension can be treated independently and no further modifications to the methods are necessary. The x and y positions of each tag are then predicted by separate parameterizations of Equations 11 and 12.

Barriers to tag motion, such as coastlines or thermal fronts, tend to preclude analytical solutions to the equations of motion but are relatively easy to handle in a numerical context. The position of each tag is updated at regular intervals by numerically integrating the velocity equations. When the tag encounters the barrier, it reacts according to some prescribed behavior pattern. Subsequently, the numerical integration proceeds as described earlier. The appropriate behavior prescription depends on the type of barrier and the nature of the tagged object. Some common choices include reflecting, absorbing, and sticking barriers—but the suite of possibilities is endless.

Extensions to incorporate variable recovery rates

This part of the article addresses the possibility of modifying trajectory-based estimators to accommodate situations where the diffusive displacements are not negligible and the recovery rates are not homogeneous in time and space. The matter essentially condenses to determining the theoretical probability density of the position of recovered tags so that an appropriate maximum-likelihood solution can be formulated. The predictor is not an issue because it is, by nature, independent of the recovery rates.

The simplest way to approach the problem is in terms of sampling strategies. Each recovered tag can be thought of as a nonrandom selection from the underlying probability density of the tagged population. Tags recovered at any specific location x are therefore misrepresented in the sample by a factor $P_R(x)$, which is the probability of recovering a tag at x. For example, if the underlying probability density for the in situ positions of the tags was the normal distribution with variance βT, then the adjusted objective function would be

$$\sum_i \frac{(x_i - \hat{x}_i)^2}{P_R[x_i] T_i}. \tag{20}$$

The correction factor P_R can be any measure of the relative likelihood of a tag arriving at any given position. In principle, it can be expressed as a function of the parameters of the underlying models of the population dynamics and estimated as part of the overall parameter search. Towards this end, it is important to recognize that the only relevant recovery processes are those that vary among tags. For example, if faster individuals were not significantly more likely to shed their tags than were slower individuals, then tag-shedding would be irrelevant to trajectory-based models regardless of its magnitude. The same would not be true for abundance-based models because tag-shedding would help to determine the total number of recoveries. Thus it should be possible to adjust trajectory-based estimators so that there are fewer parameters than those required by abundance-based estimators.

Bayliff and Rothschild (1974) developed an approximate procedure for adjusting Jones's estimators by the level of fishing effort at the time and place each tag was recovered. A similar procedure could be applied to the estimators developed here by appropriately weighting the objective function. Equation 19, for example, could be modified to read

$$\sum_i \frac{(x_i - \hat{x}_i)^2}{f_i T_i},$$

where f_i is the observed fishing effort in the area where the i'th tag was recovered. Variations in other factors, such as natural mortality and nonreporting, could be treated analogously. Bayliff and Rothschild pointed out that this method of adjustment does not account for the effort in the time-area strata encountered by the tag before its recovery. Even so, such approximate adjustments may be sufficient to reduce the degree of bias to negligible levels. If true, the estimation problem associated with Equation 20 would be simplified considerably. This is an area that deserves more attention, perhaps on an ad hoc basis via simulations.

In some situations it may be possible to circumvent the problem of inhomogeneous recovery rates by dividing the feasible domain into regions where the recovery rates are approximately constant and by estimating the advection field in each region separately. Tags of fish at liberty long enough to have strayed into any of the other regions would be excluded from the analysis.

Acknowledgments

I thank J. Cramer, V. Restrepo, and S. Turner for their comments on an early draft prepared as a working document for the ICES working group on long-term management measures (Miami, 1994). The present manuscript benefitted immeasurably from the comments of M. Prager, W. Richards, and two anonymous reviewers.

Literature cited

Bayliff, W. H., and B. J. Rothschild.
1974. Migrations of yellowfin tuna tagged off the southern coast of Mexico in 1960 and 1969. Inter. Am. Trop. Tuna Comm. Bull. 16:3–43.

Beverton, R. J. H., and S. J. Holt.
1957. On the dynamics of exploited fish populations. Fish. Invest. Minist. Agric., Fish. Food (G.B.), Series II, 533 p.

Bloomfield, P., and W. L. Steiger.
1983. Least absolute deviations: theory, applications, and algorithms. Birkhauser Verlag, Boston, MA.

De Verdiere, A. C.
1983. Lagrangian eddy statistics from surface drifters in the eastern North Atlantic. J. Mar. Res. 41:375–398.

Deriso, R. B., R. G. Punsly, and W. H. Bayliff.
1991. A Markov movement model of yellowfin tuna in the Eastern Pacific Ocean and some analyses for international management. Fish. Res. (Amst.) 11:375–395.

Fiadero, M. E., and G. Veronis.
1984. Obtaining velocities from tracer distributions. J. Phys. Oceanog. 14:1734–1746.

Figueroa, H. A., and D. B. Olson.
1989. Lagrangian statistics in the South Atlantic as derived from SOS and FGGE drifters. J. Mar. Res. 47:525–546.

Hampton, J.
1991. Estimation of southern bluefin tuna *Thunnus maccoyi* natural mortality and movement rates from tagging experiments. Fish. Bull. 89:591–610.

Hilborn, R.
1990. Determination of fish movements patterns from tag recoveries using maximum likelihood estimators. Can. J. Fish. Aquat. Sci. 47:636–643.

Hines, A. H., and T. G. Wolcott.
1990. Blue crab movement and feeding measured by ultrasonic telemetry. Bull. Mar. Sci. 46(1):246.

Jones, R.
1959. A method of analysis of some tagged haddock returns. J. Cons. 25(1):58–72.
1976. The use of marking data in fish population analysis. FAO Fish. Tech. Pap. 153, 42 p.

Kleiber, P., and A. Fonteneau.
1994. Assessment of skipjack fishery interaction in the eastern tropical Atlantic using tagging data. FAO Fish. Tech. Pap. 336/1:94–107.

Krauss, W., and C. W. Boning.
1987. Lagrangian properties of eddy fields in the northern North Atlantic as deduced from satellite-tracked drifter buoys. J. Mar. Res. 45:259–291.

Okubo, A.
1980. Diffusion and ecological problems: mathematical models. Springer-Verlag. New York, NY, 254 p.

Porch, C. E.
1993. A numerical study of larval retention in the Southern Straits of Florida. Ph.D. diss. Univ. Miami, Coral Gables, FL, 245 p.

Porch, C. E., V. R. Restrepo, S. C. Turner, and G. P. Scott.
In press. Virtual population analyses of Atlantic bluefin tuna incorporating movement and tagging data. ICCAT Col. Vol. Sci. Doc. (Working document SCRS/94/73).

Poulain, P., and P. P. Niiler.
1989. Statistical Analysis of the Surface Circulation in the California Current System Using Satellite-Tracked Drifters. J. Phys. Oceanog. 19(10):1588–1603.

Press, W. H., B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling.
1986. Numerical recipes. Cambridge Univ. Press. Cambridge, MA, 818 p.

Quinn, T. P.
1988. Estimated swimming speeds of migrating adult sockeye salmon. Can. J. Zoo. 66(10):2160–2163

Rousseeuw, P. J.
1984. Least median of squares regression. J. Am. Stat. Assoc. 79:871–880.

Saila, S. B., and J. M. Flowers.
1969. Toward a generalized model of fish migration. Trans. Am. Fish. Soc. 98:582–588.

Salvado, C. A. M.
1994. Discrete population field theory for tag analysis and fishery modeling. FAO Fish. Tech. Pap. 336/1:122–136.
Schulz, U., and R. Berg.
1992. Movements of ultrasonically tagged brown trout (*Salmo trutta* L.) in Lake Constance. J. Fish Biol. 40:909–917.
Schweigert, J. F., and C. J. Schwarz.
1993. Estimating migration rates for Pacific herring (*Clupea pallasi*) using tag recovery data. Can. J. Fish. Aquat. Sci. 50:1530–1540.
Sibert, J. R.
1984. A two-fishery tag attrition model for the analysis of mortality, recruitment and fishery interaction. Tech. Rep. 13, Tuna and Billfish Assessment Progr., South Pacific Comm., Noumea, New Caledonia, 27 p.
Sibert, J. R., and D. A. Fournier.
1994. Evaluation of advection-diffusion equations for estimation of movement patterns from tag recapture data. FAO Fish. Tech. Pap. 336/1:108–121.
Skellam, J. G.
1951. Random dispersal in theoretical populations. Biometrika 38:196–218.
Stachowitsch, M.
1979. Movement, activity pattern, and role of a hermit crab population in a sublittoral epifaunal community. J. Exp. Mar. Biol. Ecol. 39:135–150.
Taylor, G. I.
1921. Diffusion by continuous movements. Proc. London Math. Soc. 20:196–212.
Wunsch, C.
1989. Tracer inverse problems. *In* D. L. T. Anderson and J. Willebrand (eds.), Oceanic circulation models: combining data and dynamics, p. 1–77. Kluwer Academic Publs., Dordrecht.

Appendix

Define a function G such that

$$\int_{x_0}^{x} \frac{dx}{u[x]} = G[x] - G[x_0],$$

where x is the tag's current position in space, x_0 is the tag's initial position, and u is the velocity.

Then

$$\frac{d}{dx}\int_{x_0}^{x} \frac{dx}{u[x]} = \frac{dG[x]}{dx}.$$

From the fundamental theorem of integral calculus

$$\frac{1}{u[x]}\frac{dx}{dx} - \frac{1}{u[x_0]}\frac{dx_0}{dx} = \frac{dG[x]}{dx},$$

which reduces to

$$u[x] = \frac{1}{\frac{dG}{dx}}.$$