Use of Monte Carlo Methods as an Aid to the Design of Tuna Tagging Experiments

John R. Sibert and Peter J. Bills

JIMAR, Pelagic Fisheries Research Program University of Hawaii, Honolulu, Hawaii

Abstract. Data from tag and recapture experiments are important sources of information about the dynamics of fish populations. Such experiments are expensive to conduct, so it is important to know, at least in theory, whether the data collected will be informative about the processes of interest. Models based on diffusion-advection-reaction equations are used to generate hypothetical distributions of tagged fish under simple assumptions about tuna movement, mortality and fishing, and under different tag release strategies. These distributions are treated as expected values of a random variable and used to generate samples of recaptured fish. The recapture samples are analyzed and the estimated model parameters are compared with their "true" values to assess the accuracy of the procedure.

Introduction

Knowledge of rates of fish population growth, movement and death is central to the rational management of fisheries. This information is used in models of population dynamics to estimate whether a fish population can sustain exploitation and whether the effect of fisheries operating on remote fishing grounds may adversely impact fisheries for the same species on local fishing grounds ("fisheries interaction"). The release and recapture of tagged fish is an important tool used by fishery biologists to estimate the parameters of population dynamics models. Tagging experiments for tunas typically involve the release of many thousands of tagged fish from vessels operating on the high seas and can be very expensive. It is therefore useful to attempt to know in advance whether such expensive experiments are feasible. The purpose of this study is to assess the feasibility of mounting a tuna tagging experiment for the yellowfin and bigeye tuna fisheries in the Exclusive Economic Zone of the State of Hawaii. We will attempt to answer the following questions: (1) how many fish must be tagged? (2) where are the best release sites? (3) can population dynamics parameters be accurately estimated from the expected numbers of recaptures?

Considerable progress has been made in recent years in the application of diffusion-advection-reaction models to describe the population dynamics of tunas. These models have been extended to create maximum likelihood procedures for estimating population dynamics parameters from tagging data (Sibert *et al.* 1996). We have used these models in Monte Carlo simulations to attempt to address the questions posed above (Bills and Sibert, 1997).

Methods

The aggregate density of tagged tuna (N) at a point is assumed to satisfy the following partial differential equation

$$\frac{\partial N}{\partial t} + \frac{\partial}{\partial x}uN + \frac{\partial}{\partial y}vN = \frac{\partial}{\partial x}\left(D\frac{\partial N}{\partial x}\right) + \frac{\partial}{\partial y}\left(D\frac{\partial N}{\partial y}\right) - N\sum_f q_f E_f - MN \tag{1}$$

where E_f is the observed fishing effort of fleet f, q_f is a fleet specific proportionality constant or catchability coefficient, and M is the rate of "natural mortality" (mortality not attributable to fishing or migration).

The diffusive (D) and advective (u, v) terms in this equation are intended to describe the net movement of fish due both to their behavior and to components of passive transport resulting from their occupation of a moving fluid. Tunas are very strong swimmers and we generally assume that the largest component of their movement is behavioral. Therefore we regard the diffusive terms as descriptions of random components and the advective terms as descriptions of directed components of behavior.

Boundary and initial conditions are required to complete the specification of equation (2). Either closed or open boundary conditions can be used. Closed conditions should be specified around islands and along continental coastlines. Closed conditions along external model boundaries should also be specified if the entire habitat is modeled or if there is strong reason to believe the tags do not venture outside the model region. In this case, a reflection condition on N is used together with an impermeability condition on the normal directed component of movement. Thus

$$\frac{\partial N}{\partial x} = 0 \quad \text{and} \quad u = 0$$

on eastern and western closed boundaries, and

$$\frac{\partial N}{\partial y} = 0 \quad \text{and} \quad v = 0$$

on northern and southern closed boundaries.

These conditions can be used to ensure that tag numbers are conserved in the numerical model when both natural and fishing mortality are zero.

Open boundary conditions should be specified along external boundaries if a subset of the known habitat is modeled and there is good reason to believe that fish leaving the model domain do not return. In this case the advective flux of tags across a given boundary is assumed to be that of tags nearby; for example,

$$\left.\frac{\partial}{\partial x} uN\right|_b = \left.\frac{\partial}{\partial x} uN\right|_i$$

where b denotes a point on the external boundary and i denotes an interior point next to the boundary. The diffusive flux of tags near external boundaries is assumed to be zero; for example,

$$\left.\frac{\partial}{\partial x}\left(D\frac{\partial N}{\partial x}\right)\right|_i = 0$$

where i again denotes an interior point close to the external boundary.

The initial condition used for N is

$$N(x,y,0) = \begin{cases} \sum_c \tilde{N}(x_c, y_c, 0; c) & \text{at release sites,} \\ 0 & \text{at other sites,} \end{cases}$$

where $\tilde{N}(x_c, y_c, 0; c)$ is the number of tags released at point (x_c, y_c) in tag cohort c at time 0. In the case of a cohort release s at a subsequent time t_s and position (x_s, y_s), the density of tags at that point is assigned to be the sum of its present value and the cohort contribution, i.e.

$$N(x_s, y_s, t_s) \leftarrow N(x_s, y_s, t_s) + \tilde{N}(x_s, y_s, t_s; s)\,.$$

Equation (2) is solved by finite difference approximation on a 25×30 grid with a spatial resolution of 60×60 nautical miles (Figure 1) and a time step of $\frac{1}{8}$ months. The alternating direction implicit (ADI) scheme (Press *et al.*, 1992) is used to solve the finite difference approximation of (2). The ADI approach uses two implicit sweeps of the model domain, the first in the x direction over the first half time step and the second in the y direction over the second half time step. The method is unconditionally stable, meaning that local instabilities are not amplified by the solution process. Further details of this application of the ADI method can be found in Sibert and Fournier (1994).

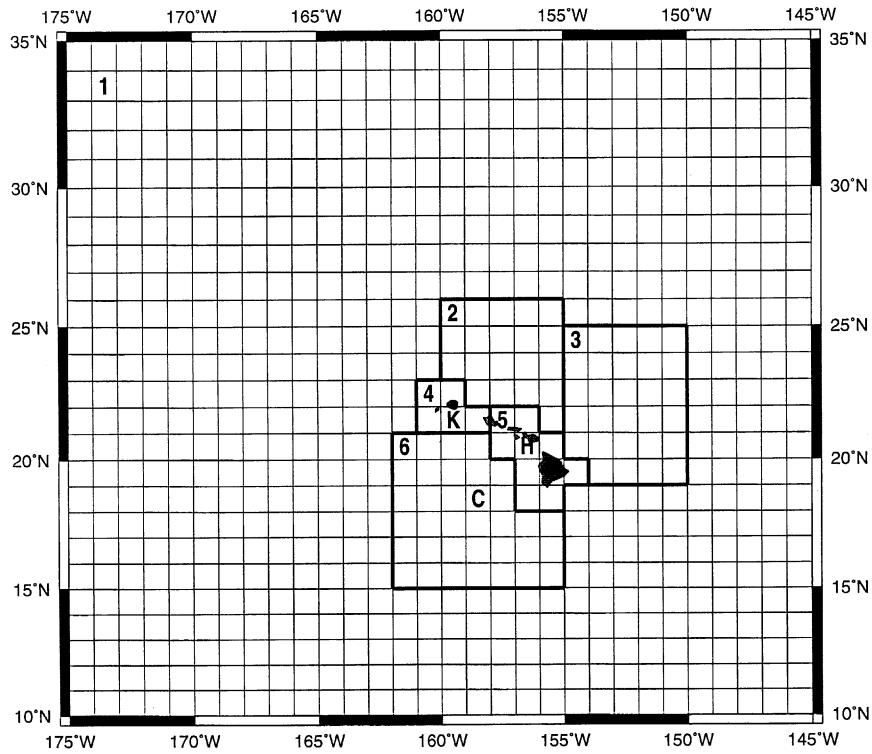

Figure 1. Regionalization of the Hawaiian model using 6 regions. K denotes the Kauai release site, H and C the Hawaii and Cross release sites.

If n denotes the entry time level and $(n+1)$ the exit level, the time derivative in (2) is differenced as

$$\left.\frac{\partial N}{\partial t}\right|_{i,j}^{n+1} \simeq \frac{N_{i,j}^{n+1} - N_{i,j}^{n}}{\Delta t}.$$

Subscripts i and j specify the spatial location of a grid point as being $(i\Delta x, j\Delta y)$; superscript n specifies time $n\Delta t$.

Upwind differencing is used to approximate the advective terms and takes the form

$$\left.\frac{\partial}{\partial x} uN\right|_{i,j}^{n} \simeq \begin{cases} \frac{u_{i,j}\, N_{i,j}^{n} - u_{i-1,j}\, N_{i-1,j}^{n}}{\Delta x}, & u_{i,j} > 0 \\ \frac{u_{i+1,j}\, N_{i+1,j}^{n} - u_{i,j}\, N_{i,j}^{n}}{\Delta x}, & u_{i,j} < 0 \end{cases},$$

$$\left.\frac{\partial}{\partial y} vN\right|_{i,j}^{n} \simeq \begin{cases} \frac{v_{i,j}\, N_{i,j}^{n} - v_{i,j-1}\, N_{i,j-1}^{n}}{\Delta y}, & v_{i,j} > 0 \\ \frac{v_{i,j+1}\, N_{i,j+1}^{n} - v_{i,j}\, N_{i,j}^{n}}{\Delta y}, & v_{i,j} < 0 \end{cases}.$$

Upwind differencing is robust but contributes numerical dispersion (O'Brien, 1986) to the solution field that may confound the estimation of the actual random movement parameters from real tagging data. In spite of this feature, upwind differencing is often used since it has the advantage of not introducing negative tag densities. Upwind differencing may also be used in the early stages of parameter estimation when parameter values are either guessed or inferred from observation.

The diffusive terms are approximated by

$$\begin{aligned} \left.\frac{\partial}{\partial x}\left(D\frac{\partial N}{\partial x}\right)\right|_{i,j}^{n} \simeq & \\ & N_{i-1,j}^{n}\,\frac{D_{i-1,j} + D_{i,j}}{2(\Delta x)^2} \\ & -N_{i,j}^{n}\,\frac{D_{i-1,j} + 2D_{i,j} + D_{i+1,j}}{2(\Delta x)^2} \\ & +N_{i+1,j}^{n}\,\frac{D_{i,j} + D_{i+1,j}}{2(\Delta x)^2}, \\ \left.\frac{\partial}{\partial y}\left(D\frac{\partial N}{\partial y}\right)\right|_{i,j}^{n} \simeq & \\ & N_{i,j-1}^{n}\,\frac{D_{i,j} + D_{i,j-1}}{2(\Delta y)^2} \\ & -N_{i,j,}^{n}\,\frac{D_{i,j-1} + 2D_{i,j} + D_{i,j+1}}{2(\Delta y)^2} \\ & +N_{i,j+1}^{n}\,\frac{D_{i,j+1} + D_{i,j}}{2(\Delta y)^2}. \end{aligned}$$

The numerical implementation of closed boundary conditions for upwind and centered-space schemes differs only for the directed movement parameters. To illustrate, consider the one-dimensional grid point lattice of Figure 2. In the x direction, the reflection conditions on N at each boundary have the numerical counterparts:

$$N_0 = N_1 \quad \text{and} \quad N_{I+1} = N_I\,.$$

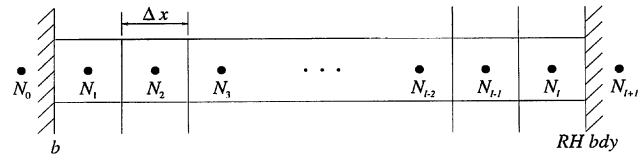

Figure 2. Grid setup showing grid point labels and closed boundaries.

If upwinding is used, the impermeability condition is differenced as

$$\begin{cases} u_0 = 0 \quad \text{and} \quad u_I = 0 \quad \text{when} \quad u < 0\,, \\ u_1 = 0 \quad \text{and} \quad u_{I+1} = 0 \quad \text{when} \quad u > 0 \end{cases}$$

and guarantees tag conservation in the absence of mortality. If the centered-space scheme is used, the impermeability condition at both ends is differenced as

$$u_0 = -u_1 \quad \text{and} \quad u_{I+1} = -u_I$$

and is tag conservative from one time level to the next if there is no natural or fishing mortality.

Now suppose the closed boundaries in Figure 2 are replaced by open boundaries. At the open left hand boundary b, a second order convergent approximation for the quantity uN at position $i = 0$ is used:

$$u_0 N_0 = 2u_1 N_1 - u_2 N_2$$

and gives rise to a first order convergent approximation for advective flux at the open boundary. If diffusive flux in the x direction is assumed to be zero at $i = 1$, the implied boundary condition is

$$\begin{aligned} N_0(D_0 + D_1) \quad = \quad & N_1(D_0 + 2D_1 + D_2) \\ & -N_2(D_1 + D_2)\,. \end{aligned}$$

The initial conditions used are $N_{i,j}^{0} = 0$ everywhere except where there is a release of tags, when N is set accordingly. For tag releases that occur after time 0, tag density at the corresponding positions is reassigned as the sum of the current and release values.

The observed distribution of fishing effort for 1991-92 was used with different tag release strategies in the time dependent solution of equation (2) to generate monthly fields of tag recapture density. Predicted densities are interpreted as expected values of a Poisson random variable. Random samples were drawn from a Poisson distribution to generate replicate sets of simulated tag recapture data. Each of these recapture data sets was analyzed by a maximum likelihood procedure to estimate the parameters of equation (2). These estimated parameters were then compared statistically with the "true" values used in the simulation.

Solution of equation (2) requires specification of all parameters at each of the 750 computational elements. Direct statistical estimation of so many parameters

from tagging data is probably not feasible. An alternative parameterization was applied. The model domain was divided into 6 regions within which the three movement parameters (u, v, D) are assumed to be constant (Figure 1). Seasons are then defined as periods of time during which the movement parameters in each region are constant. The introduction of regions and seasons reduces the number of parameters to be estimated, enables variation of tuna behavior through the year, and defines a "movement pattern".

The movement pattern used in all simulations is shown in Figure 3. This pattern implements several simple assumptions about tuna movement. Fish move southward in Autumn in expectation of unfavorable Winter conditions and northward in Spring in expectation of favorable conditions; random movement is more intense in Winter and less intense in Summer. This movement pattern requires $3 \times 6 \times 4 = 72$ movement parameters to describe the annual movement cycle, much less than the 2250 required to describe movement for a single time step of the solution of equation (2). We do not necessarily endorse these assumptions as an accurate description of tuna behavior in Hawaii. Rather these assumptions can be viewed as a plausible model to be used in generating hypothetical distributions of tag recaptures. If a tagging experiment can be designed that will accurately estimate the parameters of this pattern, we assume that the parameters of alternative movement hypotheses can also be accurately estimated. An accurate description of tuna movement in Hawaii will await the results of a real tagging experiment.

Let $[R_{i,j}]$ be a matrix containing the region number for each model element indexed by (i, j) and let $[S^n]$ be a vector containing the season for each time step

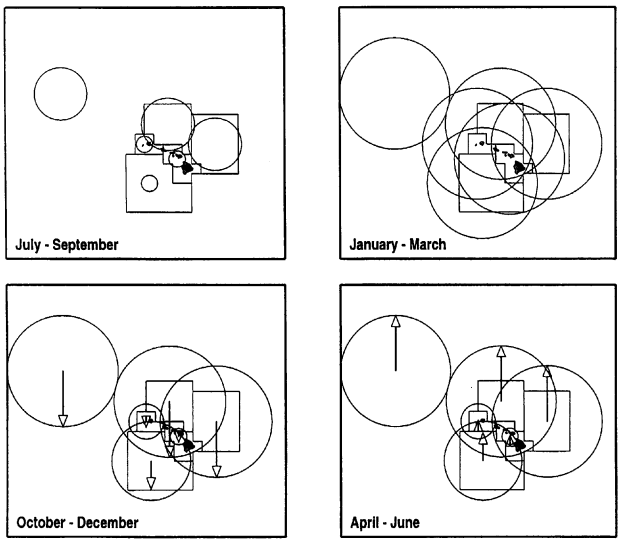

Figure 3. Four season movement pattern used in Hawaii tuna movement simulations.

indexed by n; $[R_{i,j}]$ maps the model domain into specific regions and $[S^n]$ maps calendar time to seasons. The reparameterization of the model can then be expressed as

$$[u^n_{i,j}] = \mathcal{U}_{[R_{i,j}][S^n]}, \quad [v^n_{i,j}] = \mathcal{V}_{[R_{i,j}][S^n]} \quad \text{and}$$

$$[D^n_{i,j}] = \mathcal{D}_{[R_{i,j}][S^n]},$$

where $\mathcal{U}, \mathcal{V}$ and $\mathcal{D}$ are 6×4 matrices of parameters to be estimated.

Six recognizable fishing fleets land tunas in Hawaii, and each is distinguished in the model by a characteristic catchability coefficient q_f. There are three subdivisons of the longline fleet, SSLL (targeting swordfish), TTLL (targeting tunas),and MMLL (switching between swordfish and tunas); a single troll fleet, HIRT, and two subdivisions of the handline fleet, HIIH (inshore) and HIOH (offshore). The catchability coefficients used in all model runs are shown in Table 1.

Table 1. Fleet names and corresponding catchability estimates.

Fleet	q_f
SSLL	0.000000959
TTLL	0.000000670
MMLL	0.000001889
HIRT	0.00002958
HIIH	0.00008916
HIOH	0.001300

The predicted number of tags returned during 1 month is given by

$$\widehat{C}^n_{i,j;f} = q_f E^n_{i,j;f} N^n_{i,j} \tag{2}$$

where N satisfies equation (2) at time level n.

Observed numbers of tag returns $C^n_{i,j;f}$ are related to predicted numbers of returns $\widehat{C}^n_{i,j;f}$ by the Poisson likelihood function

$$L(\mathcal{U}, \mathcal{V}, \mathcal{D}, q_f, M, C^n_{i,j;f}) = \prod_{i,j,n;f} \left[\frac{\widehat{C}^n_{i,j;f}{}^{C^n_{i,j;f}} e^{-\widehat{C}^n_{i,j;f}}}{C^n_{i,j;f}!} \right]. \tag{3}$$

This function assumes that the predicted number of tag returns in each grid cell during one month is the expected value of a Poisson random variable. The Poisson distribution is selected as it is appropriate for the observation of a rare event such as the radioactive decay of an atom (Feller, 1968) or the recapture and return of a tagged tuna.

Maximum likelihood parameter estimates are obtained by finding the values of the parameters that maximize equation (4). The maximization is accomplished by

minimizing the negative logarithm of (4) using a quasi-Newton numerical function minimizer which in turn depends on the gradient of partial derivative computed using adjoint functions (Griewank and Corliss, 1991):

$$\min_{\mathcal{U},\mathcal{V},\mathcal{D},q_f,M} \quad -\log L(\mathcal{U},\mathcal{V},\mathcal{D},q_f,M,C^n_{i,j;f}) .$$

To assist the minimizer, bounds are placed on the values of the parameters $\mathcal{U},\mathcal{V},\mathcal{D},q$ and M. Typical bounds used in the simulations are $|\mathcal{U},\mathcal{V}| < 100$, $0 < \mathcal{D} < 100000$, $0 < q < 0.1$ and $0 < M < 0.5$; parameters $\mathcal{D},q$ and M are positive and upper bounds are set to be high so as to avoid restricting the minimiser (Bard, 1974).

We evaluate the parameter estimates by examining their variability and accuracy. Variability is simply measured by the standard deviation of the estimates. Accuracy is evaluated by computing the *standardized bias* which we define as

$$\text{Standardized bias} = \frac{\text{Estimate} - \text{True value}}{\text{Standard deviation of the estimate}} . \tag{4}$$

The standardized bias statistic has unit standard deviation enabling comparison of estimates of parameters measured in different units. If the distribution of the bias estimates is normal, approximately 95% lie within two standard deviations of zero. A bias greater than 2 (or less than -2), therefore, implies a statistically significant bias.

Results

The effect of releasing tagged fish at different sites was tested. Four tests were conducted in each of which a total of 6000 tags were released. The first test was a release of 1500 tags per quarter at the release site (region 4), the second a similar release at ″

(region 5), the third at (region 6) and the fourth a release of 500 tags per quarter at each site. Figure 4 summarizes the results of 50 simulations conducted for each test. The estimate for random movement (Nm^2/mo) is plotted on the vertical axis; the natural logarithm of the average number of tag recaptures is plotted on the horizontal axis. The legend in the bottom right hand graph gives the "true" random movement value corresponding to the six regions of the model. The labeled horizontal lines on the graphs correspond to these values and provide a visual check on the accuracy of the estimates. The mean of the estimates is plotted as a black diamond. Error bars extending two standard deviations appear as vectors; the lower error bar is truncated at the x axis if calculated to extend below it—the parameter estimation is contrained to positive values. The variability in the estimates decreases as the number of recaptures increases. Estimates are least variable in those regions where the tag releases occurred.

Figure 4 shows that the accuracy of the regional movement parameters depends on the number of tags recaptured in each region. The number of recaptures can usually be increased by increasing the number of fish tagged. The decrease in variability of the estimates as the number of releases is increased can be seen in Figure 5.

Figure 6 shows the dependency of mortality components on numbers of recaptures. Catchability estimates for each fleet depend on the number of tagged tunas recaptured by that fleet. Natural mortality is a global parameter and does not depend on the total number of recaptures.

Figure 7 shows the standardized bias of the natural mortality and catchability parameter estimates and Figure 8 shows the standardized bias of the movement parameter estimates. Generally the median of the standardized bias is near zero indicating that most parameter estimates are reasonably accurate. There are a number of "outlying" individual estimates that are well outside of a normally distributed random variable, but zero bias is always between the first and third quartiles of the distributions.

Fishers and fishery managers are often concerned about the growth of one segment of the fishing fleet and its impact on other segments. The concern is that catch may decline in one area as the fishing intensifies in another area. This issue is sometimes known as "fishery interaction" and is an important motivation for tagging studies. Table 2 shows the change in recapture rates in different fleets if certain other fleets are not active. Clearly there is measurable fishery interaction in these simulations.

Table 2. Recapture percentages predicted by experiments in which fleets are selectively removed. Subtotals may not sum to the total percentage given because of rounding errors.

SSLL	TTLL	MMLL	HIRT	HIIH	HIOH	$\sum$%
0.08	1.42	2.16	1.69	1.53	1.15	8.04
0.07	1.45	2.13	1.67	1.52	–	6.84
–	–	–	1.68	1.57	1.20	4.44
–	–	–	1.69	1.58	–	3.27
–	–	–	1.73	–	–	1.73
–	–	–	–	1.62	–	1.62
–	–	–	–	–	1.22	1.22
0.09	1.43	2.18	–	–	–	3.71

Discussion and Conclusions

The most obvious conclusion of this analysis is that the accuracy of the parameter estimates depends on the number of recaptured fish. This result is hardly

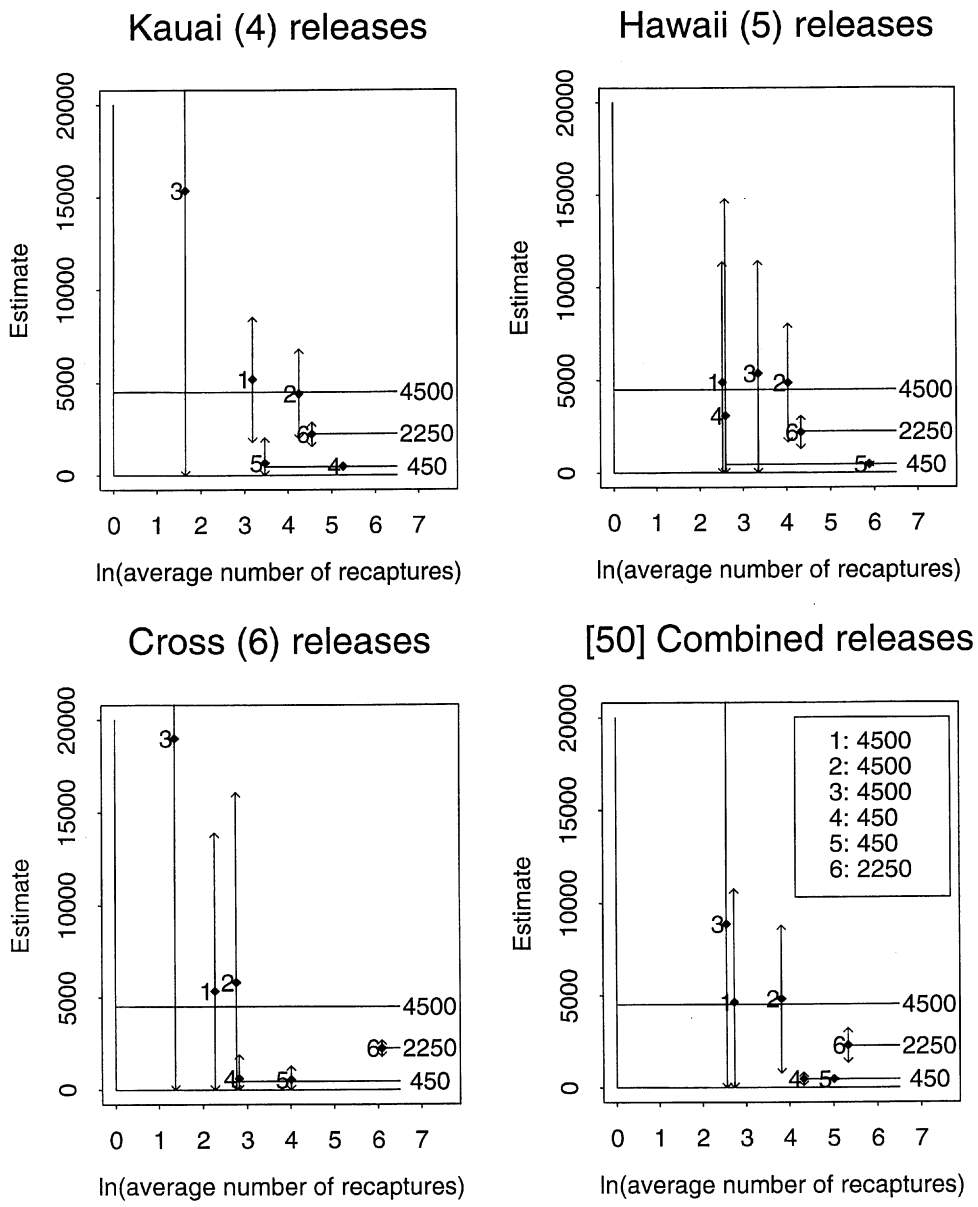

Figure 4. Random movement estimates from four tests (see text) each of 50 simulations. The legend in the bottom right plot lists "true" values of random movement for each region; the labelled horizontal lines correspond to these values. The mean of the estimates is plotted as a black diamond; error bars extending two standard deviations appear as vectors. The Kauai release site is in region 4, Hawaii in region 5 and Cross in region 6.

surprising since tag recaptures are the data on which the estimation is based.

A multiple release site strategy appears to be preferable to a single release site strategy. Releasing tags at multiple sites has the effect of distributing tags more widely within the recapture area thus increasing the probability of recapture in different regions. The relative advantages of different release sites and the relative accuracy of regional parameters will both depend on the actual movement pattern that causes fish to be recaptured in different regions. The advantage of multiple release sites for tuna tagging experiments was also noted by Bertignac (1996).

The mortality parameters are often difficult to estimate from fisheries data. The results presented here show that these parameters can be accurately estimated from tag recapture data.

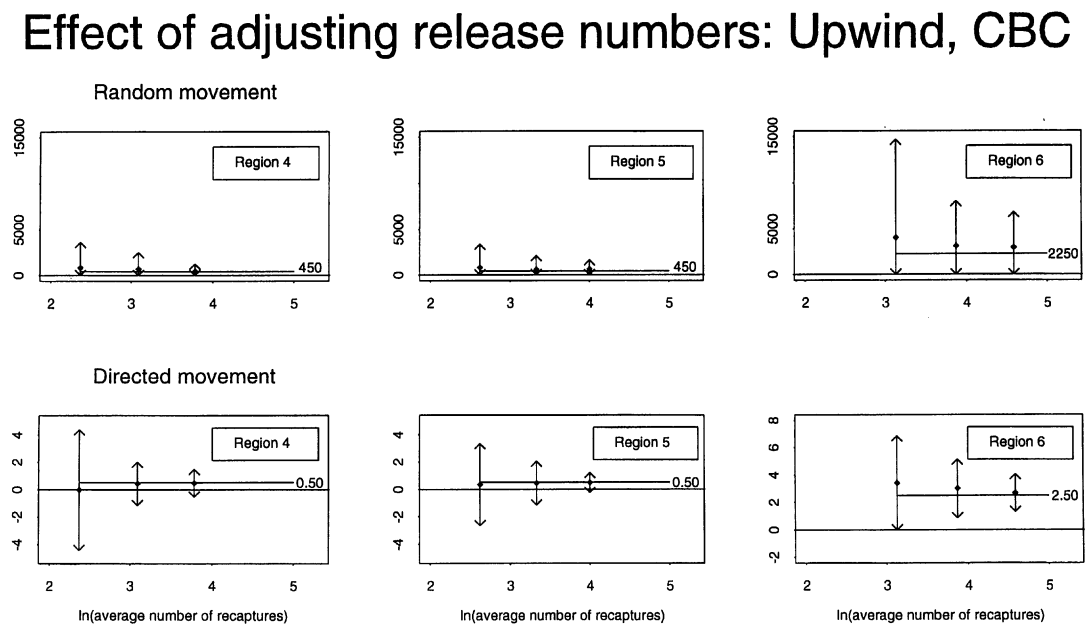

Figure 5. Mean estimates and error bars for random and directed movement using upwind closed boundary conditions with different release strategies. Results are given for Season 4, Regions 4, 5 and 6, using release rates increasing (L to R in each plot) from 250/qtr to 500/qtr to 1000/qtr. True values are marked.

Variability is a prevailing problem in fisheries research. These simulations show that, even under the controlled conditions of a numerical simulation, the possibility of extreme events is significant. To the extent that these numerical experiments can be viewed as "replications" of a tagging experiment, the possibility of a single experiment producing wildly inaccurate parameter estimates cannot be eliminated. Nevertheless, the standardized bias plots suggest that reasonable results are to be expected 99% of the time.

The use of upwind differencing is often discouraged (O'Brien, 1986) because it introduces numerical dispersion and inaccurately approximates the solution of equation (2) for given (u, v, D). The simulation and estimation procedures use the same numerical approximations and are therefore consistent. Thus we are able to obtain reasonable estimates of the input parameters from the solution field. Furthermore, there is no *a priori* theory that requires tuna to move in strict accordance to the diffusion-advection model and we do not necessarily assume that they do. Rather, we are attempting to develop a quantitative framework to express the mobility of fish populations. In this sense, the numerical approximation, or more accurately the computer code, is the model. With that consideration in mind, the numerical dispersion introduced by upwind differencing is not particularly relevant.

References

Bard, Y., *Nonlinear parameter estimation*, Academic Press, New York, 1974.

Bertignac, M. , A simulation model of tagging experiments of yellowfin in the Western Indian Ocean, in *Proceedings of the Second FAO Expert Consultation on Interactions of Pacific Ocean Tuna Fisheries; January 23-32, 1995, Shimizu, Japan*, edited by R. S. Shomura, J. Majkowski, and R. F. Harman, in press.

Bills, P. J., and J. R. Sibert, Design of tag-recapture experiments for estimating yellowfin tuna stock dynamics, mortality and fishery interaction, Joint Institute of Marine and Atmospheric Research Technical Report, in press.

Feller, W., *An Introduction to Probability Theory and its Applications*, John Wiley & Sons, New York, 1968.

Griewank, A., and Corliss, G. F.,*Automatic Differentiation of Algorithms: Theory, Practice and Application*, SIAM, Philadelphia, 1991.

O'Brien, J. J., The hyperbolic problem, in *Advanced Physical Oceanographic Numerical Modelling*, edited by James J. O'Brien, NATO ASI series, Series C, Mathematical and Physical Sciences, vol. 186, D. Reidel Publishing Company, pp. 165-186, 1968.

Press, W. H., S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in C: The Art of Scientific Computing*, 2nd ed., Cambridge University Press, 1992.

Sibert, J. R., and D.A. Fournier. Evaluation of advection-diffusion equations for estimation of movement patterns from tag recapture data, in *Proceedings of the First FAO Expert Consultation on Interactions of Pacific Ocean Tuna Fisheries; 3-11 December 1991, Noumea, New Caledonia, Vol. 1:* Summary report and papers on interaction, edited by R. S. Shomura, J. Majkowski, and S. Langi, FAO Fish. Tech. Pap. 336/1:108-121, 1994.

Sibert, J. R., J. Hampton, and D. A. Fournier, Skipjack movement and fishery interaction, in *Proceedings of the Second FAO Expert Consultation on Interactions of Paci*, edited by R. S. Shomura, J. Majkowski, and R. F. Harman, in press.

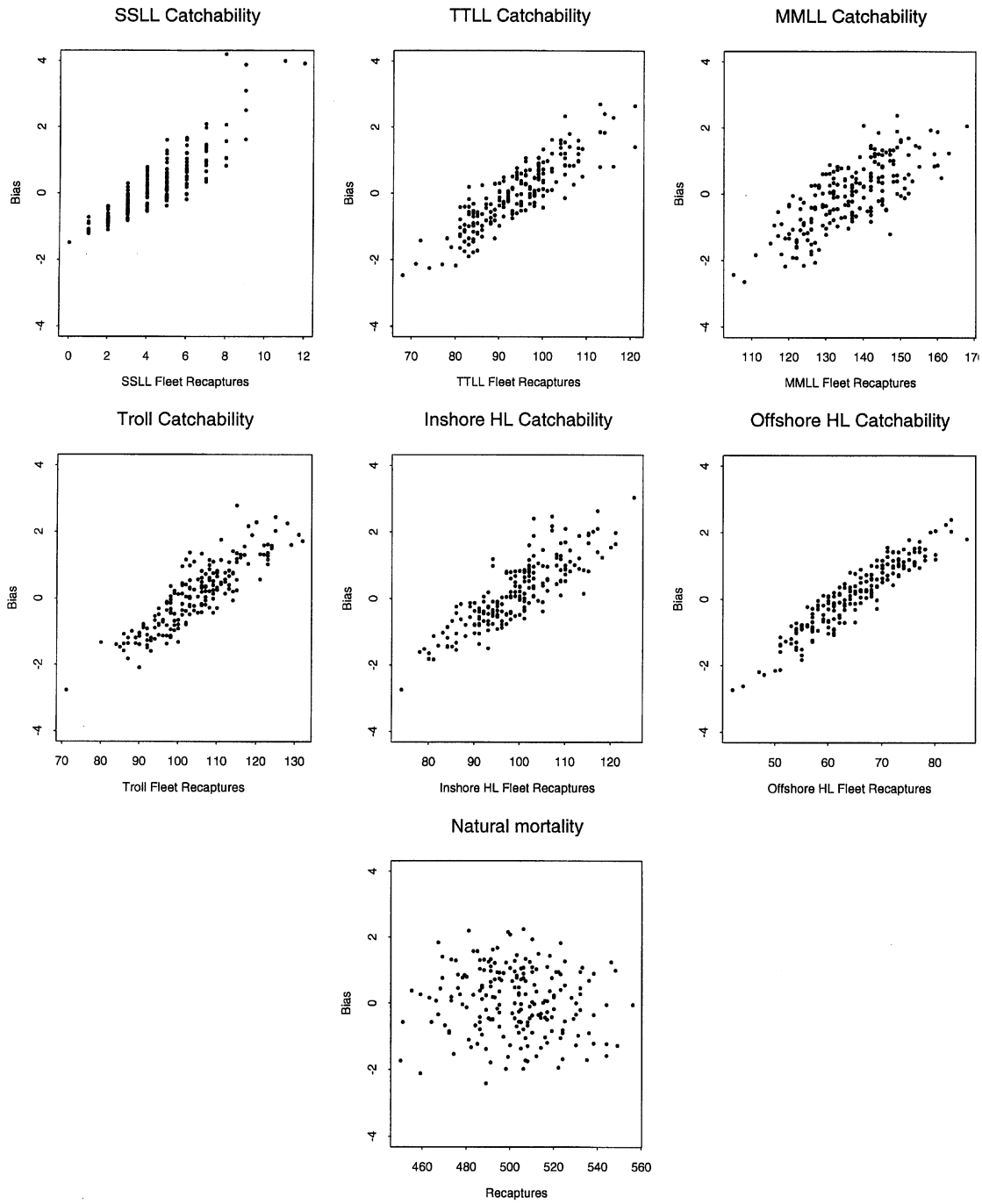

Figure 6. Scatter plots for catchability and natural mortality bias versus recaptures for 200 simulations of the multiple-site release strategy.

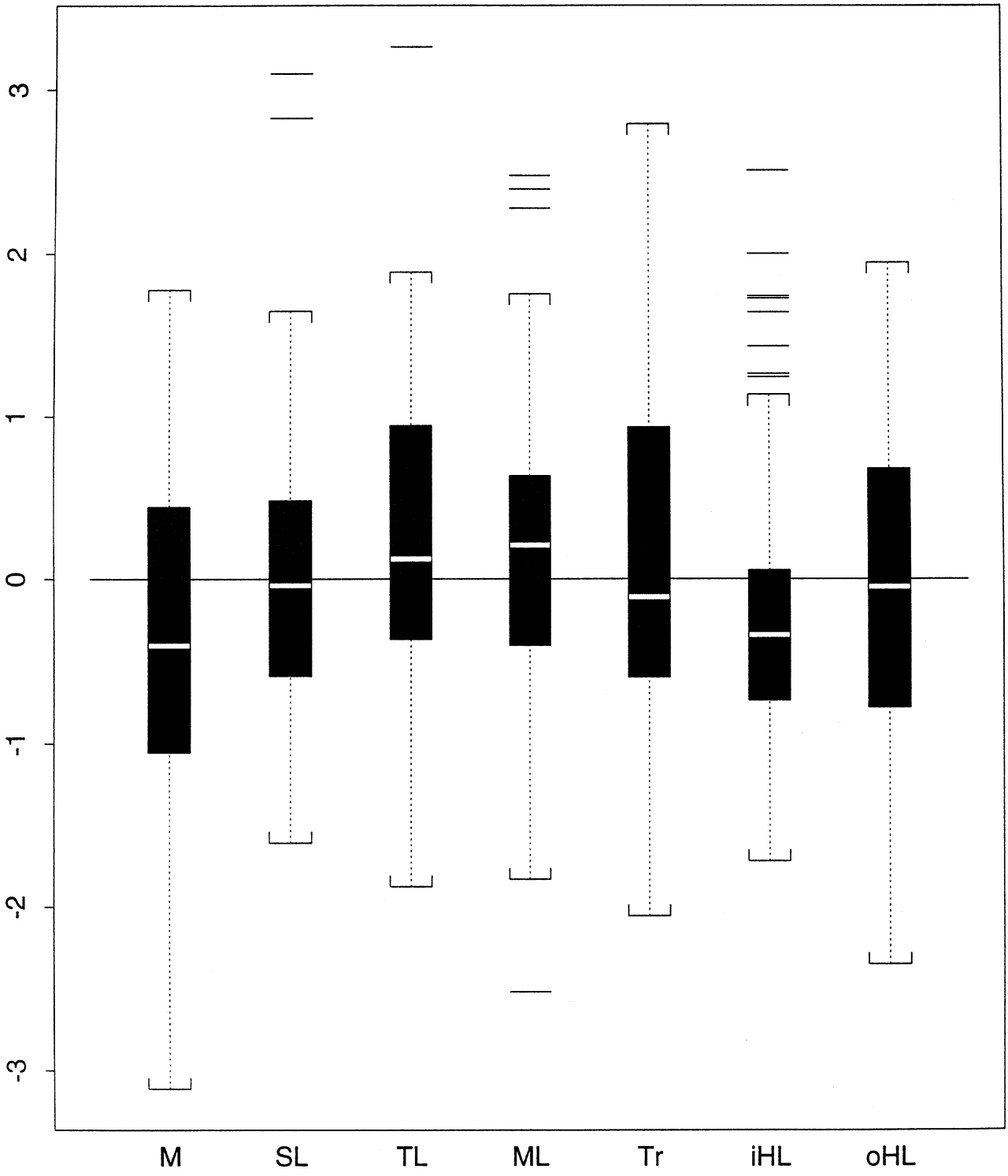

Figure 7. Distribution of standardized bias for natural mortality M and catchabilities based on 50 simulations of the upwind closed boundary scheme (SL=Swordfish LL, ..., oHL=offshore Handline). The black rectangle represents the difference between the first and third quartiles. The median is indicated by the white bar. The dotted lines leaving the top and bottom of each box extend to the extreme values of the estimates or to a distance of 1.5× the inter quartile distance, whichever is greater. The horizontal bars beyond the brackets are consider to be outliers.

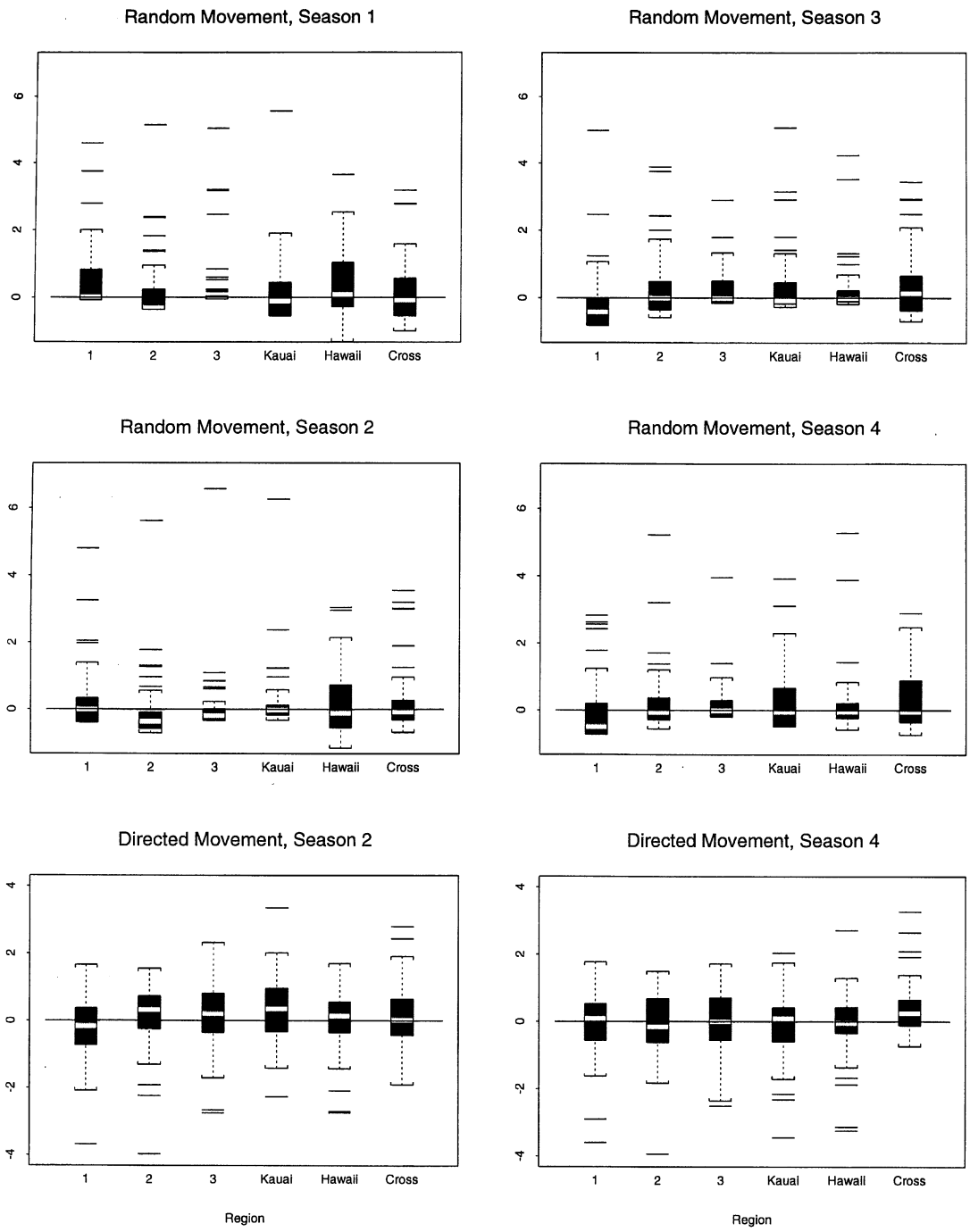

Figure 8. Distribution of standardized bias for the movement parameters based on 50 simulations of the upwind closed boundary scheme (Kauai=Region 4, etc.)