American Fisheries Society Symposium 3:149–162, 1988

Null Hypotheses, Models, and Statistical Designs in the Study of Larval Transport

DAVID R. COLBY

U.S. National Marine Fisheries Service
Southeast Fisheries Center, Beaufort Laboratory
Beaufort, North Carolina 28516, USA

Abstract.—In larval transport investigations, the broad range of relevant spatial and temporal scales, the imprecision of measurements of organism abundance, and the potentially complex motions of water masses are particular concerns in developing research designs. These concerns underscore a need both for collaboration between biologists and physicists, and for rapid adoption of emerging technologies for measuring relevant processes at appropriate scales in the field. Many branches of statistics have application to larval transport investigations. Requirements for accurate description imply a need for greater emphasis on statistical estimation procedures and attention to sources of bias, and accordingly less emphasis on statistical hypothesis testing, especially with observational data. Strong null hypotheses should be advanced that are consistent with common experience and reflect current knowledge. They then provide an invaluable framework against which to compare empirical data. In larval transport research, advection–diffusion models, physical models, explicit conceptual models, and stochastic models all can serve as null hypotheses. Random-walk, Markov-chain models of the movements and expected vertical distribution of larvae in the water column have merit for studies of vertical migration as a transport mechanism. They also can be used to model the accumulation of larvae along shorelines and other boundaries and may help explain layering of oceanic plankton. Spatial and temporal scales, target and sampling populations, independent replication, and sampling gear bias are among the concepts that require attention in the design of studies of the flux of larvae through an inlet, or of a particular transport mechanism. Because, in larval transport studies, a variable of interest may change in response to many external factors, careful examination of patterns of covariation may provide more insight than will a focus on average values. New statistical approaches, such as the bootstrap, offer promising alternatives to traditional statistical methodologies, and they underscore the increasing application of computers to all phases of larval transport research, from research design and data acquisition to data analysis and stochastic modeling.

One conclusion drawn at a recent North Atlantic Treaty Organization Advanced Research Institute on mechanisms of migration by fishes was that "lack of statistical rigor may be one of the greatest weaknesses in previous work" (McCleave et al. 1984b). This lack is understandable, given the potential spatiotemporal complexity of fish migration, the range of scales involved, the relative imprecision and unknown accuracy of techniques for measuring fish abundance, the relatively low resolution of many environmental data sets, and the difficult logistics and high costs associated with gathering such data. The biophysical nature of fish migration and larval transport points up the value of collaboration between biologists and hydrographers (Crawford and Carey 1985) and the desirability of rapid application of emerging technologies for measuring relevant processes at appropriate spatiotemporal scales.

Fish migration research is not the only area within ecology that has been criticized for a lack of statistical rigor, nor is the problem simply nonapplication of statistical concepts to the design, analysis, and interpretation of research. Problems frequently have arisen, though, because design concepts have been misunderstood and, consequently, analytical procedures have been misapplied (Underwood 1981; Hurlbert 1984).

In what follows, I discuss certain statistical concepts as they relate to investigation of larval transport through inlets. I will address these concepts primarily in regard to fishes that are spawned on the continental shelf off the southeastern United States and move landward through inlets into estuaries (Weinstein 1981; Miller et al. 1984). This context was chosen because it is one with which I am familiar, and because Oregon Inlet, connecting the Atlantic Ocean with Pamlico Sound, North Carolina, has been proposed for modification by installation of large jetties, which is of immediate interest because of the potential effect of the jetties on larval transport.

The difficulty of research into larval transport becomes more apparent when we consider what

would be required for an adequate description of, for example, the transport of larval Atlantic menhaden *Brevoortia tyrannus* from spawning sites on the edge of the continental shelf to low-salinity waters at the head of an estuary. The temporal scales extend from the minute-by-minute changes in tidal current velocities, and perhaps in the vertical elevation of the larva in the water column, to the several months between the spawning of the first egg and the arrival of the last larva at the estuarine nursery ground. The spatial scales extend from the millimeters or centimeters relevant to a larva's response to boundaries, such as the water surface or ocean floor, to the 10s of kilometers required in describing the larva's transport route. Further, a description must take into account that the water mass in which the larva swims is itself moving through space in complex and difficult-to-predict ways (Harris 1980).

In order to provide an accurate, comprehensive, and easily assimilated understanding of the transport of Atlantic menhaden, then, a descriptive model should simulate a temporal sequence of synoptic "frames," like a motion picture, showing how the larvae move through time and space. One would want to be able to "zoom" up to observe the entire population as it dispersed through time, and to "zoom" down to observe larval behavior on small spatial and short temporal scales. All this implies a need for sampling synoptically and repeatedly at a range of spatial scales, and measuring not merely fish densities but a series of environmental variables as well. It seems likely that the data required to develop such a descriptive model will only be obtained in a piecemeal fashion over a long period of time, if at all. However, several researchers have demonstrated that remarkable insights into larval transport phenomena can be gained with rather limited resources if ingenious sampling designs and novel analytical approaches are used.

Hypothesis Testing versus Estimation

Platt (1964) argued that science advances most rapidly through the systematic application of an analytical inductive method he termed "strong inference." The scientist, Platt contended, should approach any problem by first devising a set of alternative hypotheses, then designing and carrying out "crucial" experiments that will permit exclusion of one or more of the hypotheses. Repeated application of these steps produces a branching logical tree and rapid progress in understanding. It is not irrelevant that he chose to illustrate the power of the method with examples from experimental molecular biology and high-energy physics.

Quinn and Dunham (1983) rejected "strong inference" as an appropriate model for investigating ecological phenomena, arguing that "postulated ecological causes or relationships can rarely be strictly disproven, although they may often be shown to be unimportant or improbable." They go on to point out that, because patterns and processes in natural communities generally result from many causal factors that act simultaneously and interactively, the objective of research is usually to evaluate the relative importance of the different causal factors, rather than to distinguish between several mutually exclusive alternative hypotheses.

At first glance, testing hypotheses through the formal methodology of statistical analysis is attractive because it appears to offer an objective basis for separating signal from background noise. However, there is the inevitable trade-off between failure to detect that a null hypothesis is false (type-II error) and rejection of the null hypothesis when it is true (type-I error). Everything else being equal, if the investigator decides to reduce the probability of type-I error by specifying a priori a lower value of alpha, there is a simultaneous decrease in the power of the test to detect "real" differences. The magnitude of a difference required for "statistical significance," other things being equal, is a function of replication; at low levels of replication, the researcher may be in a very weak position to detect treatment effects, even when they are of substantial magnitude. At the other extreme, high levels of replication may allow the researcher to detect treatment effects that are of such low magnitude that they are, for all intents and purposes, trivial.

Ross (1985) has reminded us that hypothesis testing is framed in terms of risk, the money or other resources that will be lost as a result of a given decision. As such, Ross argued, "it is in fact irrelevant to science which is concerned with making sense, not conserving resources."

Box et al. (1978) introduced the concept of statistical hypothesis testing in a particularly ingenious way. They showed how a time series can be used as a "relevant reference set" for testing a hypothesis about, in their example, the yield of a chemical following modification of an industrial process. That is, their test employs the actual measured yields of the chemical prior to modification of the process as a basis of comparison

rather than a mathematical distribution such as Student's t. Their test requires no assumptions about statistical independence of successive values or about random sampling, unlike many standard techniques that are often incorrectly applied to temporal or spatial sequences of data. Box et al. also discussed the advantage of confidence interval estimation over hypothesis testing: "Significance testing in general has been a greatly overworked procedure, and in many cases where significance statements have been made it would have been better to provide an interval within which the value of the parameter would be expected to lie."

The appeal of statistical hypothesis testing is that it leads to relatively unequivocal decisions. Yet empirical data never conform exactly to the mathematical distributions underlying parametric theory, and so tests are to a greater or lesser extent approximations. The confidence limits approach, on the other hand, reminds the researcher that he has measured some average or pattern with a certain level of imprecision. The reason why the theory associated with estimation would seem to be more relevant to larval transport research is that the taking of measurements out in nature is fundamentally a procedure to enable accurate description. Whatever control is employed (e.g., depth of a sampling device) is nominal at best and therefore evidence bearing on a hypothesis is necessarily associative or correlative in principle—and this remains true whether hypothesis-testing statistics are computed or not.

Statistical hypothesis testing is most appropriate in an experimental context wherein treatments can be randomly assigned to experimental units. In larval transport research, statistical hypothesis testing seems to be most useful in the analysis of controlled experiments comparing, for example, behavioral or physiological responses of a species to different levels of a stimulus (e.g., Kelly et al. 1982). Interpretation of statistical tests of hypotheses becomes increasingly problematical as one moves from relatively simple phenomena, in which univariate hypotheses are tested with rigidly controlled experiments and treatments are randomly assigned to experimental units, towards complex multivariate phenomena measured in a natural system, in which several uncontrolled factors may simultaneously influence several response variables and treatments are not (or can not be) randomly assigned to experimental units.

The inadequacy of statistical hypothesis testing of ecological data is well illustrated by a recent paper by Millard and Lettenmaier (1986). They described a procedure for optimizing the statistical power of monitoring programs designed to detect ecological change following an environmental perturbation such as activation of a new wastewater discharge system. They admit the ambiguity implicit in the interpretation of tests of hypotheses for before–after designs "even if the results of the ANOVA [analysis of variance] indicate a change in organism density after the occurrence of the event, the change can reasonably be attributed to the event only if it is known that no other changes occurred in the environment that might also have affected the density." They then conclude: "A better approach, therefore, is to utilize control stations." But the "control station" approach simply adds additional assumptions about spatial uniformity, aside from wastewater discharge, between the control and potentially perturbed stations. Thus, a strong element of faith remains within the hypothesis test, and the skeptic can question the purpose of computing exact ratios of mean squares and P-values.

Perhaps ecologists have been oversold on the desirability of, and necessity for, formal statistical hypothesis testing because of the emphasis this topic usually receives in courses on applied statistics and in applied statistical textbooks (e.g., Snedecor and Cochran 1967; Winer 1971; Steel and Torrie 1980; Sokal and Rohlf 1981). It is true that experimentation both in the laboratory and in the field can provide profound insights into ecological phenomena, but it is also true that ecology is an observational science, necessarily concerned with accurate description of uncontrolled phenomena out in the natural world. It is for the latter reason that the statistical training of ecologists deserves more emphasis on the statistical theory of sampling and survey design. This branch of statistics focuses, for example, on the efficiency (in terms of costs and precision) of alternative sampling designs for estimation of parameters, and on ways of quantifying and dealing with bias from various sources (e.g., Francis 1984; Hankin 1984; Schweigert et al. 1985). Ecologists, perhaps, have not sufficiently appreciated that the validity of parameter estimates based upon sampling depend crucially upon the way in which the sample units to be actually measured are selected from among those in the population. The term "scientific sample" denotes a sample in which the measured sample units are selected with known (not necessarily equal) probability.

Multivariate analysis (Pielou 1984), time-series analysis (Bloomfield 1976), spectral analysis (Denman 1975), and circular-distribution analysis (Batschelet 1981) are other branches of applied statistics that seem particularly useful in larval transport research.

Weak versus Strong Null Hypotheses

Whether one chooses to employ statistical hypothesis testing or not, one still should formulate null hypotheses. The role of null hypotheses and so-called "null models" in community ecology has received a great deal of attention in recent years (Strong et al. 1984). Although their value for addressing certain ecological phenomena is debatable (Quinn and Dunham 1983), properly formulated null hypotheses can provide a necessary basis of reference against which empirical data may be compared (Strong 1980).

Too often, biologists have been content with weak or trivial null hypotheses that specify, for example, that mean values are equal or that a regression or correlation coefficient is zero, even when experience indicates such a null hypothesis is false, foolish, or absurd. Trivial null hypotheses of this form are intellectually unsettling if for no other reason than that they imply *nothing* is known about the phenomenon in question. To quote Ross (1985): "If knowledge were accumulating the incumbent hypothesis would hardly ever be that of no correlation or no effect. Yet this is what is tested routinely, as if all that had gone before counted for nought."

The null hypothesis for a specific case should specify the current state of knowledge of the investigator with regard to a phenomenon. For example, abundant scientific evidence and common experience confirm that larval Atlantic menhaden spawned on the continental shelf are subsequently found within North Carolina estuaries. The question then is not whether transport occurs, but rather: "Are more Atlantic menhaden transported to the estuaries than one would expect based on random dispersal?" Atlantic menhaden are highly fecund (Nelson et al. 1977), and it is conceivable that their abundance in estuaries has obscured the possibilities (1) that the vast majority of each cohort fails to reach estuaries because they disperse to inhospitable environs and subsequently die, and (2) that variation in annual cohort strength in estuaries simply reflects variation in annual patterns of circulation on the continental shelf. If we are to assess the relative effectiveness of transport, we must account for those elements of the spawn that do not attain their nursery ground, as well as those that do (Leggett 1984). This would require an extensive program to monitor the currents, eggs, and larvae on the outer continental shelf and the Gulf Stream for several months, in each of several years. In any case, it is apparent that many larval fishes are carried far beyond their normal ranges (Markle et al. 1980; Flierl and Wroblewski 1985).

Models provide the grist for specifying nontrivial null hypotheses. Fortier and Leggett (1982) have skillfully employed advection–diffusion models in larval transport research and DeAngelis and Yeh (1984) have demonstrated the utility of biased random-walk models for simulating transport. Miller et al. (1984) have described seasonal circulation patterns over the continental shelf off North Carolina. A null hypothesis appropriate for investigation of cross-shelf transport off North Carolina might, therefore, consist of a three-dimensional, biased, random-walk model that incorporates the circulation patterns described by Miller et al. (1984) for the time between spawning on the outer shelf and recruitment of larvae to the estuary. If mortality rates were incorporated into the model, the model could be used to provide location- and depth-specific estimates of larval density, which could then be compared with empirical measurements of larval density and current vectors at depth (Legendre and Demers 1984).

Scaled-down physical models also have application to larval transport research. For example, the construction of a scaled-down physical replica of Oregon Inlet permitted simulation of current patterns and larval transport through the inlet, both with and without the proposed large jetties in place (Hollyfield and Frankensteen 1980). These simulations could provide guidance in the allocation of sampling effort in time and space as well as a framework against which to compare empirical measurements of transport through the inlet.

Vertical Migration Models

By moving vertically in the water column, an organism may (1) enter an estuary or move landward or both, (2) maintain its approximate horizontal position within an estuary and avoid being flushed seaward, or (3) enhance its movement seaward (Rogers 1940; Pearcy 1962; Pearcy and Richards 1962; Graham and Davis 1971; Graham 1972; Able 1978; Weinstein et al. 1980; Fortier and Leggett 1982; Rothlisberg 1982). In some cases, the organism makes regular vertical movements in phase with the tide; in others, it attains a partic-

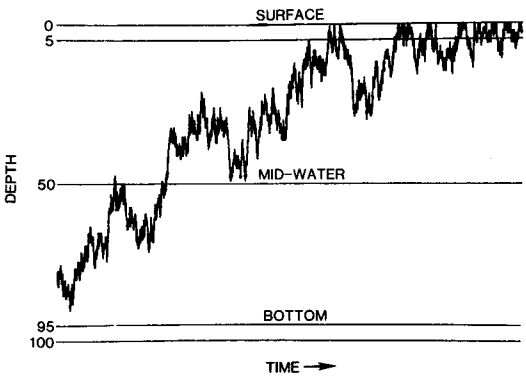

FIGURE 1.—The simulated track of an organism moving about the water column over time. The water column has an arbitrary depth of 100 units. The track represents 5,000 stochastic steps of a random walk; the organism has a 50% probability of moving either up or down, or, if it is a boundary, a 50% probability of either staying at or moving from the boundary.

ular stratum in the water column and stays there. These considerations suggest that stochastic models of vertical migration may be useful as null hypotheses in studies of vertical migration as a larval transport mechanism.

As a first model for examining the behavior of a hypothetical organism moving randomly in the vertical plane, let us consider a simple random-walk Markov chain (Parzen 1960; Hillier and Lieberman 1980; Legendre and Legendre 1983) in which the organism, with equal probability, moves either one unit up or one unit down during each time interval. Think of the water column as subdivided horizontally into several "slices" of unit thickness. On entering the uppermost slice, bounded by the surface film, the organism, during the next time interval (or step), either remains at the surface with probability p, or moves one unit downward with probability $q = 1 - p$. The corresponding specification applies to the organism if it enters the lowermost slice as well. A pseudorandom number generator on a computer was used to simulate the track of one such hypothetical organism through time for the case in which $p = q = 0.50$ (Figure 1). The transition matrix for this model, and others to be considered, will have as many rows (and columns) as there are slices making up the water column (Table 1).

One could use Monte Carlo simulations to track the organism through time and, by keeping account of how much time it spent in each slice, get a picture of its expected vertical distribution. But there is an easier, analytical, approach. Parzen (1960) showed that a Markov chain of this form will achieve a statistical equilibrium after a large number of steps (movements of the organism). One can, therefore, directly compute the stationary probabilities for each slice of the water column. For the model discussed above, where $p = q = 0.5$, the stationary probabilities are equal, meaning that, over the long term, the organism will spend an equal amount of time in each slice.

The attractiveness of the bottom and surface can be varied by varying the transition probabilities, p and q, in the top and bottom rows of the transition matrix (Table 1). Recall that q is the probability that in the next time interval the organism moves from the top or bottom slice towards midwater. If the water column is arbitrarily subdivided into 100 slices, the stationary probabilities (π) are:

$\pi_1 = \pi_{100} = (2 + 196q)^{-1}$ for the surface and bottom slices, and $\pi_2, \pi_3, \ldots, \pi_{99} = (2q)(2 + 196q)^{-1}$ for intermediate slices. So for any speci-

TABLE 1.—Transition matrix for a random-walk Markov chain in which an organism has an equal probability of moving one unit up or one unit down during each successive time interval. If it reaches the surface it will remain there during the next time interval with probability p, or move down one unit with probability $q = 1 - p$. Similarly, if it reaches the bottom it will remain there during the next time interval with probability p, or move upwards one unit with probability q.

p	q	0	0	0	•	•	•						
0.5	0	0.5	0	0	•	•	•						
0	0.5	0	0.5	0	•	•	•						
0	0	0.5	0	0.5	•	•	•						
•	•	•	•	•				•	•	•	•	•	•
•	•	•	•	•				•	•	•	•	•	•
•	•	•	•	•				•	•	•	•	•	•
					•	•	•	0.5	0	0.5	0	0	0
					•	•	•	0	0.5	0	0.5	0	0
					•	•	•	0	0	0.5	0	0.5	0
					•	•	•	0	0	0	0.5	0	0.5
					•	•	•	0	0	0	0	q	p

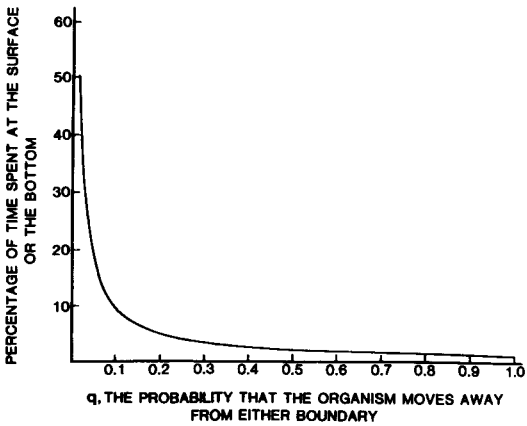

FIGURE 2.—Relationship between the expected amount of time an organism would occupy the surface or bottom "slices" of the water column and the probability that, on encountering either boundary, the organism moves toward midwater, based on a random-walk, Markov-chain model.

fied value of q, the proportion of time spent in the boundary slices will be $2(2 + 196q)^{-1}$. Only when q drops below about 0.3 does the proportion of time spent in the two boundary slices appreciably increase (Figure 2).

In the above model, the organism does not "recognize" the surface or bottom, and consequently modify its behavior, until it is within either the upper or lower one-hundredth of the water column. In nature, however, water velocity decreases as one approaches a fixed boundary (Vogel 1981), and it is worthwhile to consider a model in which the larvae take advantage of the slower velocities of the boundary layer in the lowermost portion of the water column to make greater progress against a current, or, at least, to reduce the rate at which they would otherwise be swept seaward on a falling tide. In such a model, it would be assumed that the larvae detect the lower current velocities of the boundary layer and respond by decreasing the probability that they will move upward in the water column away from the slower velocities adjacent to the bottom. A simple random-walk Markov chain to model the hypothesized boundary layer transport mechanism would have transition matrix probabilities that reflect logarithmically decreasing water velocities down through the five lowermost "slices" that make up the boundary layer, but that here, for simplicity, are held constant in the uppermost 95 slices (Table 2). Above the boundary layer, then, the organism is equally likely to move up or down, but once it enters the boundary layer its probability of moving up sharply decreases. The stationary probabilities for this model indicate that one would expect to find the hypothetical larval fish in the lowest (100th) slice about 98.8% of the time. Additional realism could be incorporated into this admittedly oversimplified model by making provision for variation in current velocities throughout the water column.

The attractiveness of this particular model for me arises from my observation of thousands of very young sockeye salmon *Oncorhynchus nerka* swimming steadily upstream through a stretch of rapids by remaining very close to the cobble bottom of a river in western Alaska. Several published studies have suggested the importance of the boundary layer to small fish likely to be swept downriver (Weinstein et al. 1980) or offshore by current (Brewer et al. 1984; Brewer and Kleppel 1986) and, therefore, lend indirect support to the boundary layer model. At very high current velocities, however, the converse situation may also apply because of bottom scouring. A correlation between catch rates of small plaice *Pleuronectes platessa* and amounts of resuspended matter in the water column led Rijnsdorp

TABLE 2.—Transition matrix for a random-walk Markov chain in which the transition probabilities for an organism reflect an approximately logarithmic decrease in water velocity in a hypothetical lower boundary layer.

0.5	0.5	0	0	0	•	•	•							
0.5	0	0.5	0	0	•	•	•							
0	0.5	0	0.5	0	•	•	•							
•	•	•	•	•				•	•	•	•	•	•	•
•	•	•	•	•				•	•	•	•	•	•	•
•	•	•	•	•				•	•	•	•	•	•	•
					•	•	•	0	0.5	0	0	0	0	0
					•	•	•	0.5	0	0.5	0	0	0	0
					•	•	•	0	0.5	0	0.5	0	0	0
					•	•	•	0	0	0.18	0	0.82	0	0
					•	•	•	0	0	0	0.07	0	0.93	0
					•	•	•	0	0	0	0	0.03	0	0.97
					•	•	•	0	0	0	0	0	0.01	0.99

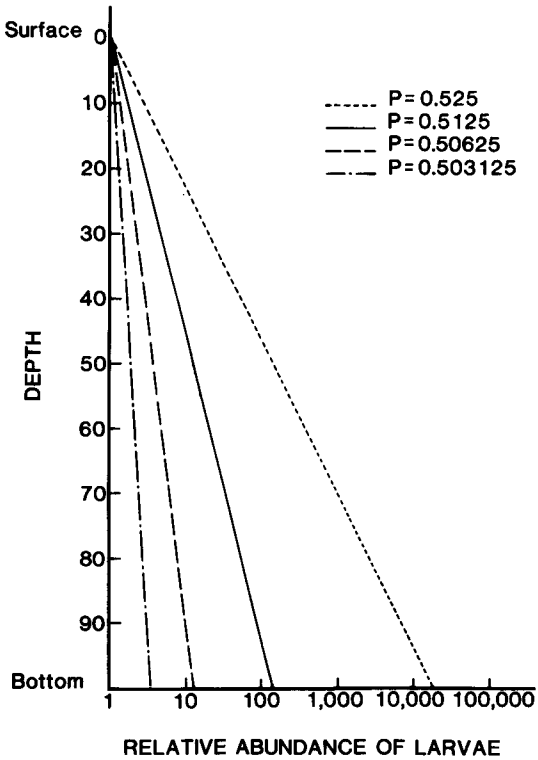

FIGURE 3.—Stationary vertical distributions of larvae in the water column that apply for different values of p; p is the probability that, in midwater, an organism will move downward, and $1.0 - p = q$ is the probability that it will move upward. The stationary probability for the ith horizontal slice of water is $(p/q)^j\ [1 - (p/q)]\ [1 - (p/q)^{100}]^{-1}$; $j = 0, 1, 2, \ldots, 99$. Relative abundance is expressed as multiples of the number of larvae at the surface.

et al. (1985) to suggest that the fish were being swept off the bottom at the time of peak tidal flows.

Finally, let us consider a model that reveals how only a very subtle vertical preference on the part of the organism implies a substantial departure from a uniform distribution. A solution presented by Parzen (1960) for a random-walk model with retaining barriers can be used to model the vertical distribution of a larva that has a greater tendency to swim upwards than downwards, or vice versa, perhaps because of light or temperature preferences, predator avoidance, or feeding. Even if the organism has only a slightly greater probability of moving downward than upward, it will spend far more time in the lower part of the water column (Figure 3). For example, if the probability of the organism moving downwards is only 0.503125 versus 0.496875 for moving upwards, the model implies that the organism should be about five times more abundant near the bottom than near the surface. Slightly increasing the probability of moving downward from 0.503125 to 0.5125 sharply increases the departure from a uniform distribution (Figure 3).

In order to use a random-walk Markov chain to model the vertical migration or distribution of a particular larval fish or invertebrate, it would be necessary to incorporate information on the depth of the water column and the organism's potential rate of movement vertically, i.e., to introduce specific scales to the model. In a more general sense, if we assume a larval fish moves forward at the rate of 1 body length/s, then we can examine the relationship between the minimum time required to move 10 m up or down in the water column, and the angle of declination of the fish's path from the horizontal (Figure 4). For example, a 20-mm larval fish tilting downward at an angle of 4° would descend 10 m in about 2 h. This sort of approach only provides a rough idea of what is possible, given certain assumptions. By way of comparison, empirical measurements of free-living, 8.1–10.0-mm yellowtail flounder *Limanda ferruginea* showed that they may move vertically as much as 25 m in 3 h (Smith et al. 1977). Furthermore, many marine invertebrate larvae can move vertically at rates of 10–50 m/h (Mileikovsky 1973; Righter 1973).

I have described these random-walk Markov-chain models in terms of the vertical dimension, with the surface and bottom or, for example, surface and pycnocline as boundaries, but they can serve in the horizontal plane as well to model phenomena such as the accumulation of organisms along a frontal system. Because, in general, the current will decrease as one approaches the lateral boundaries of an estuary or tidal creek, the boundary layer model could be conceptually rotated 90° to simulate the aggregation of larvae along shorelines.

Statistical Design Considerations

I next discuss certain statistical design concepts as they apply to measuring larval transport through coastal inlets. It is assumed that the purpose of such a study would be to either measure the flux of larvae or to test hypotheses about mechanisms employed by larvae to effect their transport.

An initial consideration is that of scales (Haury et al. 1977; Harris 1980; Gagnon and LaCroix 1982). An organism may move through a medium that is itself moving, and questions of appropriate

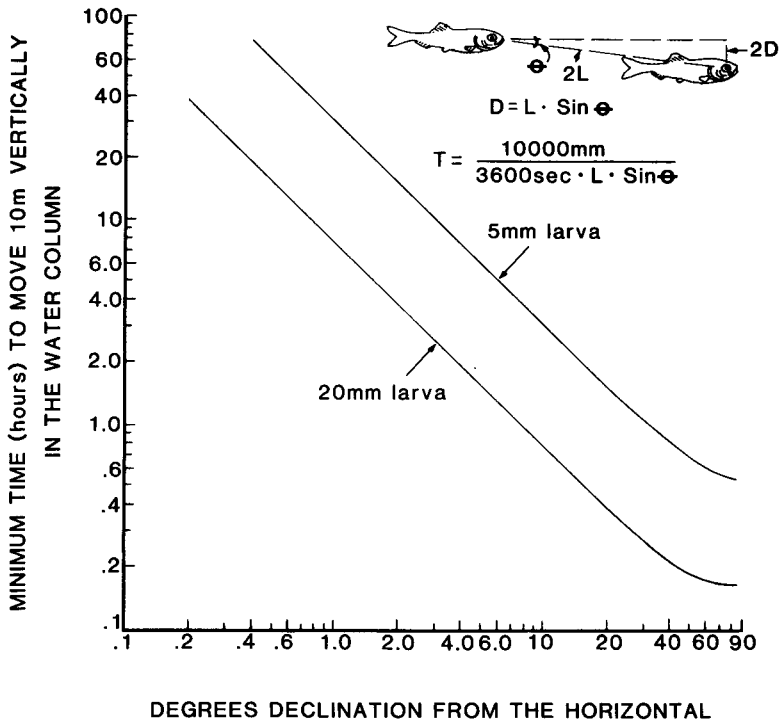

FIGURE 4.—Relationship between the rate of vertical movement (T) and the angle of declination from the horizontal for a larval fish moving forward at one body length (*L*) per second.

temporal and spatial scales for the design arise immediately. If, for example, the hypothesis involves the larva's response to physical features having scales on the order of 10s of centimeters, the design and the sampling device must provide for replicate sampling on the order of centimeters. Similarly, if the system under study involves any periodic source of variability (e.g., tide) then the design must provide for a sampling frequency at least four times the frequency of the periodic influence in order to avoid a biased and overly simplistic representation of the system (Platt and Denman 1975; Kelly 1976). If episodic, short-term events are hypothesized to have major effects on transport, the sampling period must be long enough to encompass several such events, as well as nonevent periods for comparison. If the purpose of the research is to compare flux of larvae across a cross-section of an inlet under different environmental conditions, one needs to be aware of sources of bias that may arise from flux computation methods based on the product of average velocity times average concentration (density) (see Boon 1978).

A second consideration is the care that must be exercised to ensure that the population of sample units from which a sample will be selected (sampling population), coincides with the population of sampling units for which the investigator wishes to make inferences (target population) (Cochran 1977). In larval transport research, a sampling unit might be a certain volume of water in a specified time interval. The target population might then be specified as all those volumes of water contained within some geographical boundaries during the specified time interval. However, if all of the sample units are not accessible to the sampling device, the sampling population will not coincide with the target population, and inferences about the target population will require extrapolation. The magnitude of the bias will depend upon the extent to which the excluded sample units differ in terms of the dependent variable(s) from those in the sampling population.

A third statistical concept is that of replication. The validity of many types of statistical inferences is grounded in independent replication at the design level(s) at which inferences will be made.

This means that, if inferences are to be made about the vertical distribution of a particular organism in a particular estuary during a particular time interval, samples must be taken repeatedly and independently throughout the water column within that time interval and location. If inferences are to be more general about the vertical distribution of this species in estuaries, for example, samples must be taken from more than one estuary. If one wants to make inferences about the consistency of an annual pattern, one must replicate at the level of years, etc. The more varied external environmental factors are during the period, the stronger one's inferences are; i.e., the fewer the caveats that need to be attached to the inferences.

The question of how much replication is adequate is not easily answered even though formulae for estimating sample size requirements are available (e.g., Cochran 1977; Cohen 1977; Green 1979). A conventional procedure is to estimate the variation likely to be encountered either by doing some preliminary sampling or by making calculations from existing data. In theory, this approach has much to recommend it, but, because sample variances themselves have variances, estimates of required sample size may be far from the mark. For example, the distribution of sample variances from a negative binomial distribution, often an appropriate model for positively skewed distributions of catch rates, shows that estimated sample size requirements should be viewed as rough approximations (Figure 5).

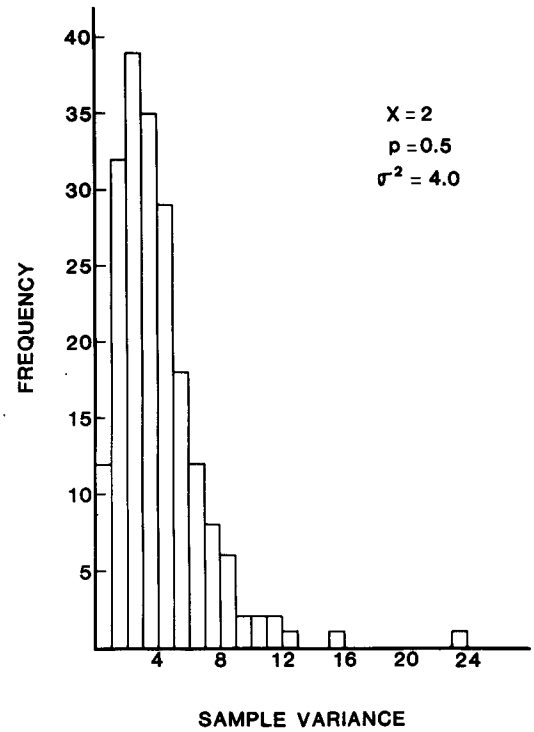

FIGURE 5.—The frequency distribution of sample variances for 200 random samples of size 10 from a negative binomial distribution; X is the Bernoulli success parameter, σ^2 is the population variance; p is the Bernoulli probability parameter.

Replication levels often are determined not by the inherent variability of the variable(s) being measured, but by constraints imposed by the availability of sampling equipment, personnel, and other resources. When this is the case, estimates of inherent variability can be used to estimate the likely precision of parameter estimates. This information then can be used to refine the design in terms of allocation of total sampling effort. At this stage, decisions can be made about what information to sacrifice and what variables to sample at higher intensity. In extreme cases, there may be such an imbalance between required precision and available resources for sampling that it makes no sense to continue beyond the initial planning phase. The patchiness of ichthyoplankton means, in general, that relatively high levels of replication will be required to estimate mean densities with even a moderate level of precision. To develop a description of the movements of young plaice over the tidal cycle in a 100 × 1,000-m area, Kuipers (1973) took "about a thousand" samples with a small trawl.

Some of these statistical considerations will now be addressed, first in the context of an investigation of larval flux through an inlet, and later in the context of examining a hypothesis about a transport mechanism.

We might begin a design for a larval flux study by deciding on the three-dimensional space that is relevant. Inasmuch as inferences are to be made about transport through the inlet as a whole, the water within the entire cross-section of the inlet must be included within the sampling population. Unless the water is homogeneous with respect to chemical and physical properties and organism densities, the efficiency of the design would be increased by subdividing the cross-section into two or more strata, portions of the cross-section that are expected to be relatively homogeneous. Note that there is no requirement that all the sampling units making up a stratum be contiguous, although this may usually be the case.

If we wish to obtain a broader picture of circulation patterns and larval densities, we would extend the sampling seaward and landward. Because larvae can attain more than intermittent (tidal) residency in the estuary only if they move from the water mass being advected in and out on the tidal cycle, we might especially wish to extend the sampling landward to where turbulent diffusion and active swimming (Fortier and Leggett 1982) permit the larvae to enter estuarine water. Further, it would be desirable to extend the sampling seaward from the inlet to include any areas that might be expected to accumulate larvae because of hydrography (Tanaka 1985).

The temporal dimension of the design will depend upon the null hypothesis as well. For example, if vertical migration is the hypothesized mechanism of transport, stratification of the design in the temporal dimension would be advantageous because one could then disproportionally allocate sampling effort to those times when the mechanism is predicted to be most operational. On the other hand, if the hypothesis involves the importance of irregular, sporadic events, the sampling must extend over a sufficiently long period to include several of these events. In the sense that the flux of larvae through an inlet is like the movement of a fluid through a transparent pipe, irregularities (bubbles in the fluid) may be especially enlightening about the rate of transport. Thus, if the sampling frame extends from seaward of the inlet to well up in the estuary, it may be feasible to track irregularities in larval densities (or cohorts, if they can be identified; see Graham and Townsend 1985) through time and space and thereby estimate rates of movement.

After stratifying the sampling frame into what seems to be relatively homogeneous sets of sample units, one is faced with decisions about sampling gears, concurrent measurements of chemical and physical variables, replication levels, and sampling frequency. All of these will depend upon particular systems, organisms, and hypotheses, so it is difficult to generalize beyond the suspicion that designs for the study of transport mechanisms will be far easier to implement than designs for measuring flux (Imberger et al. 1983). When those attempting to measure the flux of sediments, particulate matter, or nutrients in the estuary throw up their arms in dispair (Kjerfve et al. 1978), there seems no cause for optimism for those interested in accurately measuring the flux of actively swimming fishes and crustaceans.

On the other hand, if the hypothesis concerns a transport mechanism, then relatively simple designs can be very effective. In a series of papers, Fortier and Leggett (1982, 1983, 1984, 1985) demonstrated the insights to be gained from relatively simple designs that involve sampling several depth strata at frequent intervals for periods of several days. The high frequency and concurrent sampling of fish larvae, current velocities and directions, salinities, and temperature allowed Fortier and Leggett to employ spectral analysis and cross-correlation analysis to relate characteristics of the fish to characteristics of the water mass from which they were collected. They then placed these findings in the context of other information on the circulation of the St. Lawrence estuary and biology of the fish to give us some of our clearest evidence of transport mechanisms for larval fishes.

Averages and Short-Term Events

If one is sampling from a single statistical distribution, it makes sense to focus on measures of central tendency such as the arithmetic or geometric mean. But if one is dealing with a dynamic variable that continually changes in response to several external factors, it makes more sense to focus upon that variable's pattern of change and its relationship to the patterns of change of environmental factors. The spatial and temporal heterogeneity of zooplankton patches, for example was recently underscored by Omori and Hamner (1982), who reported that swarms of some zooplankters can have 1,000 times the average population density. Where that level of heterogeneity is common, the meaning of an average is elusive, and the planktologist's preoccupation with patchiness probably reflects an implicit recognition of this. In this regard, Mackas et al. (1985) pointed out that it is thus only the exceptional organism that lives out its life under "average conditions."

Several authors have argued that more attention should be concentrated on measuring transport during what may be relatively ephemeral events, such as the passage of cold fronts (Weinstein 1981; Yoder 1983; Shaw et al. 1985). It is during brief, anomalous environmental conditions that we may gain our greatest insights into transport mechanisms, either because the mechanisms only function episodically or because environmental extremes induce extreme variation in transport rates and thereby increase the chances of detecting underlying mechanisms. Because of the inherent unpredictability of such events, how-

ever, one may be forced to sample for protracted periods in order to observe enough anomalies to draw conclusions.

Additional Considerations

Ontogenetic changes in the physiology and behavior of larvae imply that the efficiency of a given sampling gear may change substantially during the period of transport from the outer continental shelf to estuarine waters. A slowly towed bongo net may accurately measure the density of larvae that are less than a week old and well up in the water column, but it may give very biased estimates of older larvae that are more concentrated in a boundary layer or are strong enough swimmers to avoid the net. If an organism's transport through an inlet is largely contingent on its acquisition of swimming competence, sampling techniques must change from the outer continental shelf to the estuary to maintain comparable capture efficiencies. As Omori and Hamner (1982) argued, too great an emphasis on standardization of sampling gear can easily lead to inaccurate measurements of plankton densities. They pointed out that the question should determine the sampling technique, not the other way around. They went on to state: "If we all sample the same way, we are constrained to examine our data from the same point of view, models will be further refined, and the data will take on a fascinating precision . . . which is mistaken for true and accurate description."

Longhurst (1984) described how new sampling instruments have changed our concepts of the pattern of biota in the upper part of the water column: "Two principal findings have emerged: first, that some of the observed horizontal patchiness was an artifact because horizontal layers do not always lie quite flat, so that their peaks and troughs may be recorded as spurious patchiness in the horizontal plane; second, and much more importantly, that the layering of biota and related variables is of a repeatability and complexity previously unsuspected."

Concluding Remarks

The theory of experimental design, sample survey design, and such parametric procedures for statistical inference as the analysis of variance, have been around for decades. However, the increasing availability of digital computers in recent years has made feasible the development of an entirely new methodology for assessing statistical error. These new methods extend, and to some extent replace, older analytical methodologies based upon assumed mathematical distributions with nonparametric approaches that utilize the power of digital computers to rapidly manipulate, resample, and calculate. These new methods both avoid assumptions about underlying distributions and provide for more comprehensive examination of the characteristics of data (Efron and Gong 1983). These "computer-intensive" methods include the bootstrap, the jackknife, cross-validation, random subsampling, and balanced repeated replication (Efron 1982; Diaconis and Efron 1983). They share in common the feature that a large number of "fake" data sets are generated by sampling with replacement from an original data set. The desired statistic (e.g., mean, median, regression coefficient, principal component coefficient, etc.) is computed for each fake data set, and the resulting distribution of these values is then used in assessing the probable error of the corresponding estimate computed from the original data. Kimura and Balsiger (1985) have provided an example of how the bootstrap can be used to improve fish survey design, and Heltshe and Forrester (1985) used the jackknife to examine the statistical properties of Brillouin's and Simpson's indices of diversity. It seems likely that these new methods will find a great deal of application in all areas of science, including marine ecology.

Acknowledgments

I thank C. W. Krouse for assistance in computer programming, J. L. Fulford for typing the manuscript, and H. R. Gordy for preparation of the figures. The manuscript was improved by suggestions from A. B. Powell, J. J. Govoni, J. W. Whitsett, G. W. Thayer, F. A. Cross, and two anonymous reviewers.

References

Able, K. W. 1978. Ichthyoplankton of the St. Lawrence estuary: composition distribution and abundance. Journal of the Fisheries Research Board of Canada. 35:1518–1531.

Batschelet, E. 1981. Circular statistics in biology. Academic Press, New York.

Bloomfield, P. 1976. Fourier analysis of time series: an introduction. Wiley, New York.

Boon, J. D., III. 1978. Suspended solids transport in a salt marsh creek—an analysis of errors. Pages 147–160 *in* B. Kjerfve, editor. Estuarine transport processes. University of South Carolina Press, Columbia.

Box, G. E. P., W. G. Hunter, and J. S. Hunter. 1978. Statistics for experimenters: an introduction to design, data analysis and model building. Wiley, New York.

Brewer, G. D, and G. S. Kleppel. 1986. Diel vertical distribution of fish larvae and their prey in near-shore waters of southern California. Marine Ecology Progress Series 27:217–226.

Brewer, G. D., G. S. Kleppel, and M. Dempsey. 1984. Apparent predation on ichthyoplankton by zooplankton and fishes in near-shore waters of southern California. Marine Biology (Berlin) 80:17–28.

Cochran, W. G. 1977. Sampling techniques, 2nd edition. Wiley, New York.

Cohen, J. 1977. Statistical power analysis for the behavioral sciences. Academic Press, New York.

Crawford, R. E., and C. G. Carey. 1985. Retention of winter flounder larvae within a Rhode Island salt pond. Estuaries 8:217–227.

DeAngelis, D. L., and G. T. Yeh. 1984. An introduction to modeling migratory behavior of fishes. Pages 445–469 *in* McCleave et al. (1984a).

Denman, K. L. 1975. Spectral analysis: a summary of the theory and techniques. Canada Fisheries and Marine Service Technical Report 539.

Diaconis, P., and B. Efron. 1983. Computer-intensive methods in statistics. Scientific American 248(5):116–130.

Efron, B. 1982. The jackknife, the bootstrap and other resampling plans. Society for Industrial and Applied Mathematics, Philadelphia.

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

Flierl, G. R., and J. S. Wroblewski. 1985. The possible influence of warm core gulf stream rings upon shelf water larval fish distribution. U.S. National Marine Fisheries Service Fishery Bulletin 83:313–330.

Fortier, L., and W. C. Leggett. 1982. Fickian transport and the dispersal of fish larvae in estuaries. Canadian Journal of Fisheries and Aquatic Sciences 39:1150–1163.

Fortier, L., and W. C. Leggett. 1983. Vertical migrations and transport of larval fish in a partially mixed estuary. Canadian Journal of Fisheries and Aquatic Sciences 40:1543–1555.

Fortier, L., and W. C. Leggett. 1984. Small-scale covariability in the abundance of fish larvae and their prey. Canadian Journal of Fisheries and Aquatic Sciences 41:502–512.

Fortier, L., and W. C. Leggett. 1985. A drift study of larval fish survival. Marine Ecology Progress Series 25:245–257.

Francis, R. I. C. C. 1984. An adaptitive strategy for stratified random trawl surveys. New Zealand Journal of Marine and Freshwater Research 18:59–71.

Gagnon, M., and G. LaCroix. 1982. The effects of tidal advection and mixing on the statistical dispersion of zooplankton. Journal of Experimental Marine Biology and Ecology 56:9–22.

Graham, J. J. 1972. Retention of larval herring within the Sheepscot estuary of Maine. U.S. National Marine Fisheries Service Fishery Bulletin 70:299–305.

Graham, J. J., and C. W. Davis. 1971. Estimates of mortality and year class strength of larval herring in western Maine, 1964–67. Rapports et Procès-Verbaux des Réunions, Conseil International pour l'Exploration de la Mer 160:147–152.

Graham, J. J., and D. W. Townsend. 1985. Mortality, growth and transport of larval Atlantic herring *Clupea harengus* in Maine coastal waters. Transactions of the American Fisheries Society 114:490–498.

Green, R. H. 1979. Sampling design and statistical methods for environmental biologists. Wiley, New York.

Hankin, D. G. 1984. Multistage sampling designs in fisheries research: applications in small streams. Canadian Journal of Fisheries and Aquatic Sciences 41:1575–1591.

Harris, G. P. 1980. Temporal and spatial scales in phytoplankton ecology. Mechanisms, methods, models, and management. Canadian Journal of Fisheries and Aquatic Sciences 37:877–900.

Haury, L. R., J. A. McGowan, and P. H. Wiebe. 1977. Patterns and processes in the time–space scales of plankton distributions. Pages 277–327 *in* J. H. Steele, editor. Spatial pattern in plankton communities. Plenum, New York.

Heltshe, H. F., and N. C. Forrester. 1985. Statistical evaluation of the jacknife estimate of diversity when using quadrat samples. Ecology 66:107–111.

Hillier, F. S., and G. J. Lieberman. 1980. Introduction to operations research, 3rd edition. Holden-Day, Oakland, California.

Hollyfield, N. W., and E. D. Frankensteen. 1980. Oregon Inlet larval transport sensitivity study. U.S. Army Corps of Engineers, Wilmington, North Carolina.

Hurlbert, S. H. 1984. Pseudoreplication and the design of ecological field experiments. Ecological Monographs 54:187–211.

Imberger, J., and seven coauthors. 1983. The influence of water motion on the distribution and transport of materials in a salt marsh estuary. Limnology and Oceanography 28:201–214.

Kelly, J. C. 1976. Sampling the sea. Pages 361–387 *in* D. H. Cushing and J. J. Walsh, editors. The ecology of the seas. Saunders, Philadelphia.

Kelly, P., S. D. Sulkin, and W. F. Van Heukelem. 1982. A dispersal model for larvae of the deep sea red crab *Geryon quinquedens* based upon behavioral regulation of vertical migration in the hatching stage. Marine Biology (Berlin) 72:35–43.

Kimura, D. K., and J. W. Balsiger. 1985. Bootstrap methods for evaluating sablefish pot index surveys. North American Journal of Fisheries Management 5:47–56.

Kjerfve, B. K., R. Dyer, and J. R. Schubel. 1978. Epilogue: where do we go from here? Pages 319–324 *in* B. Kjerfve, editor. Estuarine transport processes. University of South Carolina Press, Columbia.

Kuipers, B. 1973. On the tidal transport of young plaice (*Pleuronectes platessa*) in the Wadden Sea. Netherlands Journal of Sea Research 6:376–388.

Legendre, L., and S. Demers. 1984. Towards dynamic biological oceanography and limnology. Canadian Journal of Fisheries and Aquatic Sciences 41:2–19.

Legendre, L., and P. Legendre. 1983. Numerical ecology. Elsevier Science Publishing New York.

Leggett, W. C. 1984. Fish migrations in coastal and estuarine environments: a call for new approaches to the study of an old problem. Pages 159–178 *in* McCleave et al. (1984a).

Longhurst, A. 1984. Heterogeneity in the ocean—implications for fisheries. Rapports et Procès-Verbaux des Réunions Conseil International pour l'Exploration de la Mer 185:268–282.

Mackas, D. L., K. L. Denman, and M. R. Abbot. 1985. Plankton patchiness: biology in the physical vernacular. Bulletin of Marine Science 37:652–674.

Markle, D. F., W. B. Scott, and A. C. Kohler. 1980. New and rare records of Canadian fishes and the influence of hydrography on resident and nonresident Scotian Shelf ichthyofauna. Canadian Journal of Fisheries and Aquatic Sciences 37:49–65.

McCleave, J. D., G. P. Arnold, J. J. Dodson, and W. H. Neill, editors. 1984a. Mechanisms of migration in fishes. Plenum, New York.

McCleave, J. D., F. R. Harden Jones, W. C. Leggett, and T. G. Northcote. 1984b. Fish migration studies: future directions. Pages 545–554 *in* McCleave et al. (1984a).

Mileikovsky, S. A. 1973. Speed of active movement of pelagic larvae of marine bottom invertebrates and their ability to regulate their vertical position. Marine Biology (Berlin) 23:11–17.

Millard, S. P., and D. P. Lettenmaier. 1986. Optimal design of biological sampling programs using the analysis of variance. Estuarine, Coastal and Shelf Science 22:637–656.

Miller, J. M., J. P. Reed, and L. J. Pietrafesa. 1984. Patterns, mechanisms and approaches to the study of migration of estuarine-dependent fish larvae and juveniles. Pages 209–225 *in* McCleave et al. (1984a).

Nelson, W. R., M. C. Ingham, and W. E. Schaaf. 1977. Larval transport and year-class strength of Atlantic menhaden, *Brevoortia tyrannus*. U.S. National Marine Fisheries Service Fishery Bulletin 75:23–41.

Omori, M., and W. M. Hamner. 1982. Patchy distribution of zooplankton: population assessment and sampling problems. Marine Biology (Berlin) 72:193–200.

Parzen, E. 1960. Modern probability theory and its application. Wiley, New York.

Pearcy, W. G. 1962. Ecology of and estuarine population of winter flounder, *Pseudopleuronectes americanus* (Walbaum). Part 3. Bulletin of the Bingham Oceanographic Collection, Yale University 18(1):16–38.

Pearcy, W. G., and S. W. Richards. 1962. Distribution and ecology of fishes of the Mystic River estuary, Connecticut. Ecology 43:248–259.

Pielou, E. C. 1984. The interpretation of ecological data: a primer on classification and ordination. Wiley, New York.

Platt, J. R. 1964. Strong inference. Science (Washington, D.C.) 146:347–353.

Platt, T., and K. L. Denman. 1975. Spectral analysis in ecology. Annual Review of Ecology and Systematics 6:189–210.

Quinn, J. F., and A. E. Dunham. 1983. On hypothesis testing in ecology and evolution. American Naturalist 122:602–617.

Righter, G. 1973. Field and laboratory observations of the diurnal vertical migration of marine gastropod larvae. Netherlands Journal of Sea Research 7:126–134.

Rijnsdorp, A. D., M. van Stralen, and H. W. van der Veer. 1985. Selective tidal transport of North Sea plaice larvae *Pleuronectes platessa* in coastal nursery areas. Transactions of the American Fisheries Society 114:461–470.

Rogers, H. M. 1940. Occurrence and retention of plankton within the estuary. Journal of the Fisheries Research Board of Canada 5:164–171.

Ross, J. 1985. Misuse of statistics in social sciences. Nature (London) 318:514.

Rothlisberg, P. C. 1982. Vertical migration and its effect on dispersal of penaeid shrimp larvae in the Gulf of Carpentaria, Australia. U.S. National Marine Fisheries Service Fishery Bulletin 80:541–554.

Schweigert, J. F., C. W. Haegele, and M. Stocker. 1985. Optimizing sampling design for herring spawn surveys in the Strait of Georgia, B.C. Canadian Journal of Fisheries and Aquatic Sciences 42:1806–1814.

Shaw, R. F., W. J. Wiseman, Jr., R. E. Turner, L. J. Rouse, Jr., R. E. Condrey and F. J. Kelly, Jr. 1985. Transport of larval gulf menhaden *Brevoortia patronus* in continental shelf waters of western Louisiana: a hypothesis. Transactions of the American Fisheries Society 114:452–460.

Smith, W. G., J. D. Sibunka, and A. Wells. 1977. Diel movements of larval yellowtail flounder, *Limanda ferruginea*, determined from discrete depth sampling. U.S. National Marine Fisheries Service Fishery Bulletin 76:167–178.

Snedecor, G. W., and W. G. Cochran. 1967. Statistical methods, 6th edition. Iowa State University Press, Ames.

Sokal, R. R., and F. J. Rohlf. 1981. Biometry: the principles and practice of statistics in biological research, 2nd edition. Freeman, San Francisco.

Steel, R. G. D., and J. H. Torrie. 1980. Principles and procedures of statistics, 2nd edition. McGraw-Hill, New York.

Strong, D. R. 1980. Null hypotheses in ecology. Synthesis 43:167–171.

Strong, D. R., D. Simberloff, L. G. Abele, and A. B. Thistle, editors. 1984. Ecological communities: conceptual issues and the evidence. Princeton University Press, Princeton, New Jersey.

Tanaka, M. 1985. Factors affecting the inshore migration of pelagic larval and demersal juvenile red sea bream *Pagrus major* to a nursery ground. Transac-

tions of the American Fisheries Society 114:471–477.

Underwood, A. J. 1981. Techniques of analysis of variance in experimental marine biology and ecology. Oceanography and Marine Biology: an Annual Review 19:513–605.

Vogel, S. 1981. Life in moving fluids. Willard Grant Press, Boston.

Weinstein, M. 1981. Plankton productivity and the distribution of fishes on the southeastern U.S. Continental Shelf. Science (Washington, D.C.) 214:351–352.

Weinstein, M. P., S. L. Weiss, R. G. Hodson, and L. R. Gerry. 1980. Retention of three taxa of postlarval fishes in an intensively flushed estuary, Cape Fear River, North Carolina. U.S. National Marine Fisheries Service Fishery Bulletin 78:419–436.

Winer, B. J. 1971. Statistical principles of experimental design, 2nd edition. McGraw-Hill, New York.

Yoder, J. A. 1983. Statistical analysis of the distribution of fish eggs and larvae on the southeastern U.S. Continental Shelf with comments on oceanographic processes that may affect larval survival. Estuarine, Coastal and Shelf Science 17:637–650.