

## Replication and Treatment Strength in Whole-Lake Experiments

Stephen R. Carpenter

*Ecology*, Vol. 70, No. 2 (Apr., 1989), 453-463.

Stable URL:

<http://links.jstor.org/sici?sici=0012-9658%28198904%2970%3A2%3C453%3ARATSIW%3E2.0.CO%3B2-6>

*Ecology* is currently published by The Ecological Society of America.

---

Your use of the JSTOR archive indicates your acceptance of JSTOR's Terms and Conditions of Use, available at <http://www.jstor.org/about/terms.html>. JSTOR's Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Please contact the publisher regarding any further use of this work. Publisher contact information may be obtained at <http://www.jstor.org/journals/esa.html>.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

---

JSTOR is an independent not-for-profit organization dedicated to creating and preserving a digital archive of scholarly journals. For more information regarding JSTOR, please contact [support@jstor.org](mailto:support@jstor.org).



## REPLICATION AND TREATMENT STRENGTH IN WHOLE-LAKE EXPERIMENTS<sup>1</sup>

STEPHEN R. CARPENTER<sup>2</sup>

Center for Limnology, University of Wisconsin, Madison, Wisconsin 53706 USA

**Abstract.** Replication of large-scale experiments is desirable, but the numbers of replicates needed are not known. Costs and feasibility of ecosystem experiments depend critically on the numbers of replicates needed because of the high cost per replicate and the scarcity of experimental ecosystems. This paper examines the numbers of replicates and magnitudes of manipulation needed to detect changes in lake primary productivity resulting from piscivore manipulations. Substantial ( $\approx 10\times$  changes in the independent variable, piscivore biomass) and sustained (at least 3-5 yr) manipulations using five reference and five experimental ecosystems produced significant *t* test results in  $>80\%$  of simulated experiments. The need for substantial and sustained manipulations is consistent with published results of whole-lake experiments on nutrient inputs, chemical contaminants, and the biota.

In many cases, limited numbers of experimental systems or high costs will prevent adequate replication of ecosystem experiments. When large-scale experiments employ insufficient (e.g., 2 or 3) replicates and/or modest perturbations of the independent variate, there is great risk of erroneously accepting the hypothesis of no treatment effect. Therefore, unreplicated paired-system experiments (one reference and one experimental system) are often preferable even though classical statistics cannot be used to determine whether manipulation caused a change in the experimental system. A series of unreplicated paired-system experiments, staggered in time and performed in many locations, will provide more ecological insight than a replicated experiment in a single region. Few statistical methods pertain to large-scale ecological experiments; innovations could be very beneficial.

**Key words:** cascading trophic interactions; ecosystem experiment; experimental design; food web; lakes; press experiment; pulse experiment; replication.

### INTRODUCTION

Whole-ecosystem experiments have proven extremely valuable in ecology (Likens 1985, National Research Council 1986) but pose special statistical problems because of the difficulty of replication and the great temporal variability of ecosystems. The challenges of analyzing whole-ecosystem experiments are to determine if changes have occurred in the manipulated system and whether any changes can be attributed to the manipulation (Frost et al. 1988). There are essentially two approaches to these issues. The first requires experimental and reference (or control) ecosystems but does not require replication. A massive manipulation is employed to extensively alter the experimental system, and descriptive data are adduced to demonstrate that (1) changes in the manipulated system are much greater than those in the reference system and (2) those changes are most plausibly explained by the manipulation. This approach avoids statistical errors largely by avoiding statistics altogether (Hurlbert 1984). Such nonstatistical analyses have been used in many ecosystem-level experiments (Likens

1985); examples include Likens et al. (1970), Schindler (1977), and Carpenter et al. (1987). Intervention analyses are available to test unreplicated data for post-treatment changes in the experimental system that may have been caused by the manipulation (Box and Tiao 1975, Box et al. 1978, Stewart-Oaten et al. 1986, Frost et al. 1988, Carpenter et al. 1989). The second approach is the genuinely replicated, properly randomized experiment, which allows direct statistical tests for effects of manipulation but is rarely used at the ecosystem level (Hurlbert 1984). I view the use of multiple reference systems by Schindler et al. (1985) as an example of the replicated approach. A refinement of this approach is the "staircase" design of Walters et al. (1988), which accounts for potentially confounding time-treatment interactions.

### WHOLE-LAKE EXPERIMENTS

Whole-ecosystem experimentation originated in limnology (Likens 1985), and lakes provide one of our richest bodies of data on system-level responses. Whole-lake experiments incorporate a broad range of ecosystem phenomena, and have consequently focused our understanding of eutrophication (Schindler 1977), acidification (Schindler et al. 1985), other chemical stresses (Schindler 1988), and predator effects on ecosystem processes (Henrikson et al. 1980, Shapiro and

<sup>1</sup> Manuscript received 17 November 1987; revised 1 June 1988; accepted 23 June 1988.

<sup>2</sup> Present address: Department of Biological Sciences, University of Notre Dame, Notre Dame, Indiana 46556 USA.

Wright 1984, Carpenter et al. 1987). Whole-lake experiments can be particularly useful for determining the relative importance of conflicting mechanisms identified by experiments at smaller scales. For example, whole-lake nutrient-addition studies showed that evidence of CO<sub>2</sub> limitation derived from bottle bioassays was not germane to lake eutrophication (Schindler 1977, 1988).

Climatic factors and food web dynamics are the major sources of variability that affect whole-lake experiments (Carpenter and Kitchell 1987, Schindler 1987). Changes in fish community structure alter herbivory and thereby cause changes in primary production as large as those caused by changes in physicochemical factors (Carpenter et al. 1987, Carpenter and Kitchell 1988). Variability in food web dynamics may be the largest source of variance confounding whole-lake contaminant experiments (Schindler 1987). With the exception of the whole-lake analyses discussed below, the literature on limnetic food web dynamics will not be reviewed here. Exhaustive reviews are found in Kerfoot and Sih (1987), Carpenter (1988a), and Northcote (1988).

The extreme manipulations used to demonstrate effects in unreplicated whole-lake experiments have been criticized as ecologically unrealistic (Crowder et al. 1988). Ecosystem experimenters counterargue that only strong manipulations will reveal the response potential of lakes (Kitchell et al. 1988). Nevertheless, more subtle responses of lakes are of great theoretical and practical interest (Schindler 1987, Crowder et al. 1988), and replicated ecosystem experiments may be capable of determining such impacts. These very expensive experiments should be designed with great care.

The feasibility of replicated experiments at the ecosystem level has not been analyzed, to my knowledge. Because of the variability of lake ecosystems (Carpenter and Kitchell 1987), the number of replicates needed may be prohibitively high, or detection of effects of modest treatments may be impossible. One could estimate sample sizes for whole-lake experiments on primary production using standard statistical procedures (Sokal and Rohlf 1981), estimating the standard deviation of primary production from published data (Carpenter and Kitchell 1987). However, such calculations would not account for the joint effects of interacting ecosystem components, the time lags characteristic of actual food webs, or the effects of the manipulation on the variance of the ecosystem (Carpenter 1988b). This paper overcomes these problems by analyzing a model that combines the known variability of both primary production and fish recruitment (Carpenter and Kitchell 1987) with food web dynamics determined from successful piscivore manipulations (Carpenter et al. 1987). Monte Carlo simulations examine the degrees of replication and manipulation necessary for successful experiments on lake food webs. Because of the great interest in "biomanipulation" as

a technique for improving water quality through fishery management (Shapiro and Wright 1984), this paper emphasizes experiments in which piscivores are manipulated and responses of primary production are determined.

#### THE MODEL

The goal of this study was to estimate the power of whole-lake experiments to detect effects of piscivore manipulation on primary production. Piscivore manipulation was chosen as a convenient example of an ecosystem perturbation that has relatively large effects comparable to (and potentially derived from) those of chemical contaminants (Carpenter et al. 1987, Schindler 1987). Specifically, I sought the probability  $Q$  of rejecting the hypothesis of no effect as a function of the magnitude of piscivore manipulation and the numbers of replicates employed. I used a stochastic model that transmitted piscivore effects through the food chain to primary producers, and expressed the variance of primary productivity as the collective result of weather and food chain effects. Since piscivore manipulations affected primary production in the model, simulated experiments that failed to detect these effects led to erroneous conclusions.  $Q$  was estimated by simulating a large number of experiments and determining the proportion that had statistically significant piscivore effects. Combinations of manipulation magnitude and replicate number that yield high  $Q$  values (i.e., high probabilities of detecting manipulation effects) represent desirable experimental designs.

The model simulates a four-level food chain subject to two stochastic effects: the direct effects of variable weather on the phytoplankton, and the interannual variability of fish recruitment, which affects the phytoplankton indirectly through trophic interactions. The model was calibrated using published data on variability of primary production and fish recruitment (Carpenter and Kitchell 1987) and results of food web manipulations involving three lakes during 1984–1987 (Carpenter et al. 1987, S. Carpenter, J. Kitchell, J. Hodgson, and associates, *personal observation*). An earlier paper discussed the temporal dynamics and variance structure of a very similar model (Carpenter 1988b). Here, the model was modified to determine the outcome of experimental manipulations of piscivores. Since the model and data on which it is based have been described in detail elsewhere, only essential features are described below and in Table 1.

Primary production  $P$  varies around the mean productivity  $P^*$  at mean zooplankton biomass  $Z^*$  because of weather variability and effects of grazing.

$$P(t) = P^*[(1 + E_w) + \{c_1[Z^* - Z(t)]\}]. \quad (1)$$

The direct effect of weather on primary production,  $E_w$ , is a normally distributed random variable with variance  $V_p$ . The empirically determined parameter  $c_1$

determines the response of primary production to deviations of zooplankton biomass from the mean.

Deviations from the mean zooplankton biomass result from planktivory and weather-driven variability in primary production.

$$Z(t) = Z^* + [c_2(1 + E_w)] - g'F_1(t) - G'F_2(t). \quad (2)$$

The empirical parameter  $c_2$  converts the weather effect on primary production into an effect on zooplankton biomass. The vectors  $g'$  and  $G'$  contain coefficients that measure the contribution of each fish of a specific age to the reduction of zooplankton biomass. Planktivores include a short-lived fish (age vector  $F_1$ ) that is planktivorous throughout its life cycle and the juvenile year classes of a longer lived fish (age vector  $F_2$ ). Adults of the longer lived fish are principally piscivorous, but their diets include some zooplankton.

$F_1$  and  $F_2$  represent fish densities in each age class at the end of each year. Age-specific survivorships of short-lived fishes and juvenile long-lived fishes depend on the density  $T$  of piscivorous adults of the long-lived fish population. The probability of survival from age  $i$  to  $j$  is

$$S_{ji} = S_m[1 - (T/K)], \quad (3)$$

where  $S_m$  is maximum survivorship in the absence of predation and  $K$  is the carrying capacity of the system for adult piscivores. Recruitment of young-of-the-year piscivores depends on density of adult piscivores and a stochastic component.

$$R = R_m(1 + E_R)[1 - (T/K)]. \quad (4)$$

$E_R$  is a normally distributed random variable with variance  $V_R$  which represents the effect of interannual variability in recruitment.  $R_m$  is maximum piscivore recruitment. Recruitment of young-of-the-year of the short-lived fish follows an analogous equation. The short-lived fish is semelparous and lives for 3 yr, based on the life history of redbelly dace (*Phoxinus eos*), a common planktivore in the lakes studied by Carpenter et al. (1987). The long-lived fish is iteroparous, reaches adulthood at 4 yr, and lives a maximum of 10 yr. The long-lived fish is based on data for largemouth bass (*Micropterus salmoides*), the piscivore in the lakes studied by Carpenter et al. (1987).

Previous analyses (Carpenter 1988b) did not consider the effects of correlations among weather effects  $E_w$  on neighboring lakes or correlations among the weather effect  $E_w$  and recruitment effect  $E_R$  within a lake. Regional correlations among  $E_w$ s are common (e.g., Schindler et al. 1985).  $E_w$  and  $E_R$  are correlated, but not perfectly so because recruitment depends on weather at certain critical times of the year, while primary production integrates weather effects throughout the annual cycle. Moreover, the effects of  $E_w$  on phytoplankton are immediate, while the effects of  $E_R$  on phytoplankton have time lags that depend on fish de-

TABLE 1. Parameter values used in simulations. Vectors are listed from youngest to oldest age class.  $N(a, b)$  denotes a normally distributed random variable with mean  $a$  and standard deviation  $b$ .

Variable	Value	Units
$c_1$	0.4	dimensionless
$c_2$	0.2	dry mass, g/m <sup>2</sup>
$E_R$	$N(0, 0.5)$	dimensionless
$E_w$	$N(0, 0.2)$	dimensionless
$F_1(0)$	640, 120, 40	fish/ha
$F_2(0)$	160, 30, 10, 2, 1, 1, 0, 0, 0, 0	fish/ha
$g$	0.3, 3, 6	dry mass, g/fish
$G$	0.3, 3, 6, 1.5, 0.15, 0.15, 0.15, 0.15, 0.15, 0.15	dry mass, g/fish
$K$		
Planktivore	40 000	fish/ha
Piscivore	400	fish/ha
$P^*, P(0)$	30	(as C) g · m <sup>-2</sup> · yr <sup>-1</sup>
$R_m$		
Planktivore	2	yr <sup>-1</sup>
Piscivore	1	yr <sup>-1</sup>
$S_m$		
Planktivore	0.4, 0.5	yr <sup>-1</sup>
Piscivore	0.4, 0.45, 0.5, 0.55, 0.6, 0.65, 0.7, 0.75, 0.8	yr <sup>-1</sup>
$Z^*, Z(0)$	2.5	dry mass, g/m <sup>2</sup>

velopment rate and durations that depend on piscivore longevity (Carpenter et al. 1985, Carpenter 1988b).

In this paper, correlations were included by the following method.  $X$  and  $Y$  are random variables with correlation coefficient  $r$ , and could represent weather effects on neighboring lakes or the weather effect and recruitment effect within a lake. All  $r$  values are positive in this paper.  $\text{Var}(X)$ ,  $\text{var}(Y)$ , and  $r$  are specified by the modeler.  $X$  and  $Y$  are formed from the uncorrelated random variables  $u$ ,  $v$ , and  $w$  so that

$$X = u + w, \quad (5)$$

$$Y = v + qw \quad (6)$$

where  $q$  is the ratio of the standard deviation of  $Y$  to that of  $X$ . To calculate  $\text{var}(u)$ ,  $\text{var}(v)$ , and  $\text{var}(w)$ , note that

$$\text{var}(X) = \text{var}(u) + \text{var}(w), \quad (7)$$

$$\text{var}(Y) = \text{var}(v) + q^2\text{var}(w), \quad (8)$$

$$\begin{aligned} \text{cov}(X, Y) &= \text{cov}(u, v) + q[\text{cov}(u, w) \\ &\quad + \text{cov}(v, w) + q\{\text{var}(w)\}] \\ &= q[\text{var}(w)]. \end{aligned} \quad (9)$$

Therefore

$$\text{var}(w) = r[\text{var}(X)\text{var}(Y)]^{1/2}/q. \quad (10)$$

$\text{Var}(w)$  is obtained from Eq. 10 using information specified by the modeler, and then  $\text{var}(u)$  and  $\text{var}(v)$  can be obtained from Eqs. 7 and 8.

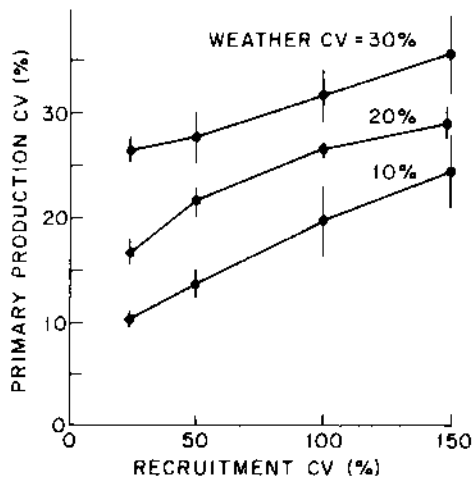


FIG. 1. Coefficient of variation (cv) of primary production vs. the cv of fish recruitment for three values of the cv of direct weather effects on phytoplankton. Data are means and standard deviations of triplicate simulations of 100 yr each.

Experiments were simulated by specifying the coefficients of variation (cvs) of weather effects and piscivore recruitment, the correlation coefficient of weather effects and piscivore recruitment, the interlake correlation coefficient of weather effects, the number of lakes to be used, and the magnitude of the piscivore manipulation to be simulated. Except where specified otherwise, cvs of weather effects and recruitment were set at 20 and 50%, respectively, near the medians of reported values (Carpenter and Kitchell 1987). Correlation coefficients are specified below. The lakes were randomly divided into two groups of equal size, 25 yr were simulated for each lake to ensure a random array of initial conditions, and manipulations were then imposed on the experimental lakes in year 26. Responses (see Results) of reference and experimental lakes were compared by a *t* test. One hundred experiments of each type were simulated, and the percentage of experiments with significant results at the  $P = .05$  level was recorded. Binomial 95% confidence intervals (Rohlf and Sokal 1981) were determined for the percentage of experiments with significant results.

Bender et al. (1984) distinguish "pulse" experiments, in which a population is manipulated at a single time, from "press" experiments in which a population is manipulated and then held at the new level for an extended period of time. Frost et al. (1988) discuss some consequences of the pulse-press dichotomy for experimentation in aquatic ecology. Both pulse and press experiments were simulated in this study. In simulated pulse experiments, a single manipulation of piscivores was performed in year 26. For pulse piscivore enhancement experiments, each lake's response was calculated as mean annual primary production in years 26, 27, 28, and 29. For pulse piscivore reduction experiments, each lake's response was calculated as mean

annual primary production in years 26 and 27. This difference between types of experiments corresponded with the different response times observed in deterministic simulations (see Results). In simulated press experiments, piscivores were manipulated in year 26 and held at the new level until year 30. In the press simulations, each lake's response was calculated as the mean primary production in years 26 through 30.

To examine the effects of correlated interlake weather effects, correlated pairs of lakes were simulated. One randomly chosen member of each pair was manipulated in year 26, and responses of primary production were calculated as above. A paired *t* test was used to test for manipulation effects. These experiments were replicated versions of the paired-system experiments commonly employed by ecosystem experimenters (Likens 1985).

## RESULTS

The coefficient of variation (cv) of primary production increased as the cv of the weather effect and the cv of fish recruitment increased (Fig. 1). In all subsequent figures, simulations used a recruitment cv of 50% and a weather cv of 20%, yielding a cv of 23% for primary production. These cvs are within the range of published values (Carpenter and Kitchell 1987).

When the cv of fish recruitment was low, the cv of primary production was equal to or less than that of the weather effect (Fig. 1) because zooplankton responses damped the response of phytoplankton to weather variability, as argued by McQueen et al. (1986) and Walters et al. (1988). Each 1% increase in the weather cv increased the cv of primary production by  $\approx 0.8\%$ . The effect of recruitment variability was much weaker. Each 1% increase in the recruitment cv increased the cv of primary production by  $\approx 0.1\%$ . This attenuation of recruitment variance through the food web is consistent with the arguments of McQueen et al. (1986).

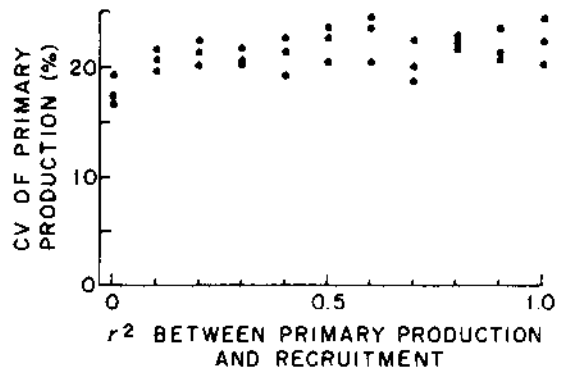


FIG. 2. cv of primary production vs. the squared correlation of primary production and fish recruitment. Results of three simulations of 100 yr each are shown for each value of  $r^2$ . For all simulations, the cv of the weather effect on primary production was 20% and the cv of fish recruitment was 50%.

When the interannual variations of primary production and recruitment were positively correlated, the cv of primary production increased (Fig. 2). However, this effect is relatively slight, raising the average cv of primary production by only 6% over the full range of the correlation coefficient. Suppose that in a given year both primary production and recruitment were above average. Since piscivores recruit as planktivores, the young-of-the-year fishes will suppress the zooplankton, thereby enhancing primary production. Conversely, when both primary production and recruitment are below average, control of zooplankton biomass by planktivores is relatively weak, and heavy grazing further reduces primary production. Thus, the effect of the correlation is to amplify the deviations of primary production from the mean, so the variance of primary production rises.

Both increases and decreases of piscivores in both pulse and press experiments were simulated for this study. Fivefold manipulations illustrate the patterns of responses to pulse manipulations in deterministic simulations (Fig. 3). The piscivore increase raised piscivore density for  $\approx 4$  yr, after which densities returned by damped oscillations to the mean. During the 4-yr period of enhanced piscivory, planktivore density was low, zooplankton biomass was high, and primary production was low, relative to the mean conditions. The piscivore reduction decreased piscivore densities for 2 yr before damped oscillations returned the population to the mean. During the 2 yr of decreased piscivory, planktivore density was relatively high, zooplankton biomass was reduced, and primary production was increased over the mean conditions. In simulated press manipulations, the magnitudes of the responses were about the same as those in the pulse manipulations, but the responses were maintained for four additional years.

Stochastic simulations illustrate typical results of press manipulations (Fig. 4). In the stochastic simulations, responses must be discerned against substantial variability similar to long-term data reported for many lakes (e.g., Goldman and de Amezaga 1984). Stochastic simulations of pulse manipulations produce changes in primary production similar to those from press manipulations, but the changes persist for a shorter time. Pulse additions of piscivores depress primary production for  $\approx 5$  yr, while pulse removals of piscivores increase primary production for 1–2 yr.

The experimentally induced changes in primary production in these simulations (Fig. 4) are smaller than those reported in some whole-lake food web experiments (Henrikson et al. 1980, Carpenter et al. 1987). These simulations evoke only a fraction of the model's maximum response of primary production to fish manipulation. Before manipulation, the simulated systems contained moderate densities of piscivores and planktivores, and 10-fold piscivore manipulations changed mean primary production (as C) by  $\approx 10$

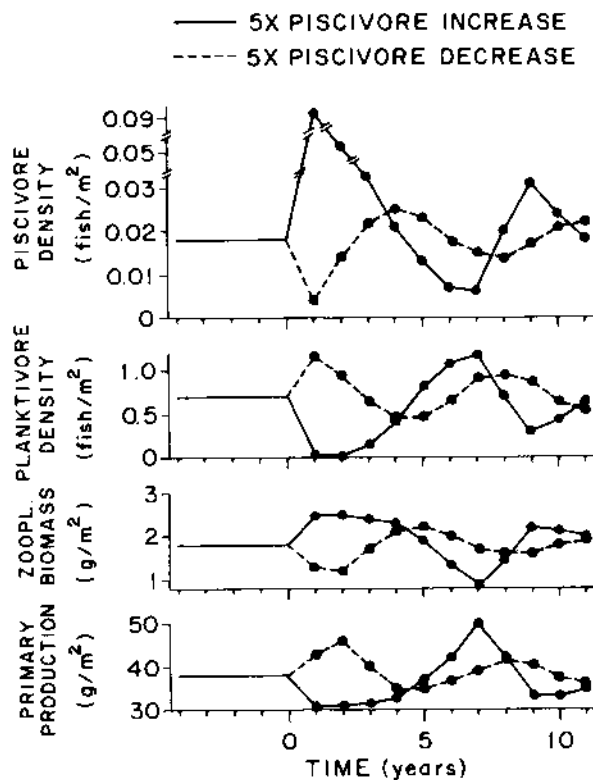


FIG. 3. Effects of pulse piscivore manipulations in deterministic simulations. Results of fivefold piscivore increase and fivefold piscivore decrease are shown. Abscissa: years relative to manipulation in year 0. Biomass is measured as dry mass per unit area.

$g \cdot m^{-2} \cdot summer^{-1}$ . If the model's premanipulation conditions are set at extremely high or extremely low piscivore densities, then comparable piscivore manipulations change primary production (as C) by  $\approx 30 g \cdot m^{-2} \cdot summer^{-1}$ . I chose to analyze initial conditions of intermediate piscivory to determine relatively conservative standards for ecosystem experiments and to examine the potential responses of a food web that may represent a typical lake.

Both replication and degree of manipulation had substantial effects on the frequency of whole-lake experiments that exhibited statistically significant effects (Fig. 5). When no manipulation was imposed ( $1 \times$ ),  $\approx 5\%$  of pulse experiments detected spurious differences between control and reference lakes, consistent with the type I error rate expected for  $t$  tests with  $P = .05$ . When only 4 or 6 lakes were used in pulse piscivore reduction experiments ( $0.1$ – $0.5 \times$ ), there was no detectable increase in the frequency of successful experiments, even when 10-fold reductions ( $0.1 \times$ ) were imposed. When 20 lakes were used, about half of the experiments detected effects of piscivore reductions ranging from  $0.1$  to  $0.33 \times$ .

In pulse experiments, effects of piscivore enhancements were detected more readily than those of pis-

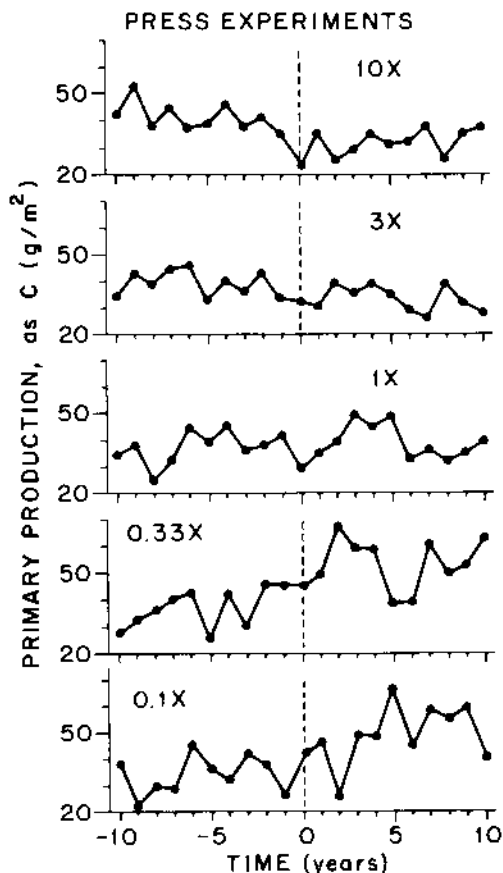


FIG. 4. Responses of primary production in stochastic simulations with 5-yr press piscivore manipulations. Abscissa: years relative to manipulation in year 0. Ordinate: primary production during summer stratification. In descending order, manipulations were: 10-fold piscivore addition ( $10\times$ ); 3-fold piscivore addition ( $3\times$ ); no manipulation ( $1\times$ ); 3-fold piscivore reduction ( $0.33\times$ ); 10-fold piscivore reduction ( $0.1\times$ ).

civore reductions (Fig. 5). However, the probability of detecting effects still depended on the magnitude of manipulation and number of replicates. To achieve a 75% probability of detecting effects, at least  $5\times$  manipulations employing at least 10 lakes were necessary.

Correlation between the weather effect on primary production and piscivore recruitment had no effect on the detection of algal responses to piscivore manipulations (Fig. 6). This observation is consistent with the weak effects of this correlation on the variability of primary production (Fig. 2). Similar trends occurred for both pulse and press experiments.

The presence of correlated interlake weather effects increased the frequency of experiments yielding statistically significant results (Fig. 7). Paired  $t$  tests have fewer degrees of freedom than unpaired designs using the same number of lakes, but achieve increased sensitivity by removing variance due to correlations of the paired systems (Box et al. 1978). Thus, when paired lakes were uncorrelated ( $r^2 = 0$ ) the paired experiments

were less successful than their unpaired counterparts because of reduced degrees of freedom (Fig. 7). However, with squared correlations as low as  $r^2 = 0.2$ , paired designs detected effects as frequently as their unpaired counterparts. Thus, even a small interlake correlation can overcome the reduction in degrees of freedom entailed by a paired-lake experiment. With interlake correlations of  $r^2 = 0.8$ , paired designs detected effects more frequently than unpaired designs. Similar results occurred for both pulse and press experiments.

Statistically significant effects were more likely with press manipulations than with pulse manipulations (Fig. 8). Press manipulations were especially advantageous in piscivore removal experiments and modest (twofold to fivefold) piscivore enhancements. Nevertheless, large sample sizes ( $\approx 10$  lakes) and relatively extensive manipulations (at least threefold) were necessary to detect effects reliably even in press experiments.

## DISCUSSION

### *Applicability of model results*

The general conclusions about replication and manipulation magnitude are (1) independent of model structure and (2) concordant with available evidence from lake ecosystems.

More complex food web models than the one used here have been employed to simulate relatively small perturbations of community structure (e.g., Bender et al. 1984, Carpenter and Kitchell 1987, Frost et al. 1988, Yodzis 1988). Many of these analyses examine perturbations around an equilibrium condition. My analysis deals with large perturbations over periods of many years, the scales of lake management. Lake systems do not appear equilibrial at these scales (Carpenter 1988a). Thus I have chosen a modeling approach that uses means as convenient reference points (which the system may occupy rarely if at all), and emphasizes the temporal variability of the system.

The food chain representation used here is simpler than the webs modeled in some other simulated experiments (e.g., Carpenter and Kitchell 1987, Yodzis 1988). The chain configuration includes the system's strong interactions, which dominate community response to perturbation (Paine 1980), and minimizes the number of coefficients that need to be estimated from field data. The chain model is similar to a more elaborate lake food web model in its responses to perturbation and probability distributions for primary production (Carpenter and Kitchell 1987). Therefore, the chain and web models would lead to similar recommendations about replication and manipulation magnitude. However, the chain model simulates experiments  $>2000$  times as fast as the web model, and so is much better suited to the repetitive analyses performed for this paper.

The model represents the mean and interannual

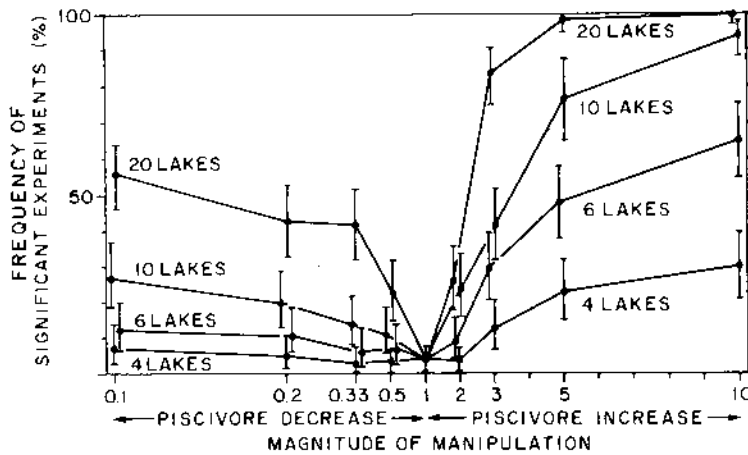


FIG. 5. Frequency of pulse experiments with significant *t* statistics ( $P < .05$ ) vs. the magnitude of piscivore manipulation for experiments with 4, 6, 10, and 20 lakes. Data are means and 95% confidence intervals.

variance of primary productivity in response to piscivore manipulations. These means and variances determine the outcomes of the simulated experiments and statistical tests. Model results are consistent with the known responses of lakes to fish perturbation in unreplicated experiments, and to the documented variances of fish recruitment and lake primary production. Since the simulated responses and variances accord with observations, the simulated statistical tests and inferences about manipulation strength and replication are reasonable.

The interannual cv of primary production in the simulated experiments ( $\approx 23 \pm 3\%$ ) is within the range of 16–30% known from many lakes (Carpenter and Kitchell 1987, 1988). The interannual cv of primary production estimated by this model matches the cv of 23% produced by a more complex food web model that included nutrient cycling, three categories of phy-

toplankton, three categories of herbivores, and invertebrate predators as well as fish (Carpenter and Kitchell 1987). The present model's interannual cv of primary production is lower than the interlake cv of 37% reported by Schindler et al. (1985). Had a higher cv been used in the model, even larger numbers of replicates and magnitudes of manipulations would have been indicated.

Mean responses to the simulated manipulations are near the middle of the range reported in whole-lake experiments. Moderate responses were achieved by manipulating systems with intermediate levels of piscivores and planktivores. Greater responses of primary production can be evoked by adding piscivores to a formerly piscivore-free system or by removing them from a piscivore-dominated system (Henrikson et al. 1980, Carpenter et al. 1987).

Available data on whole-lake food web experiments

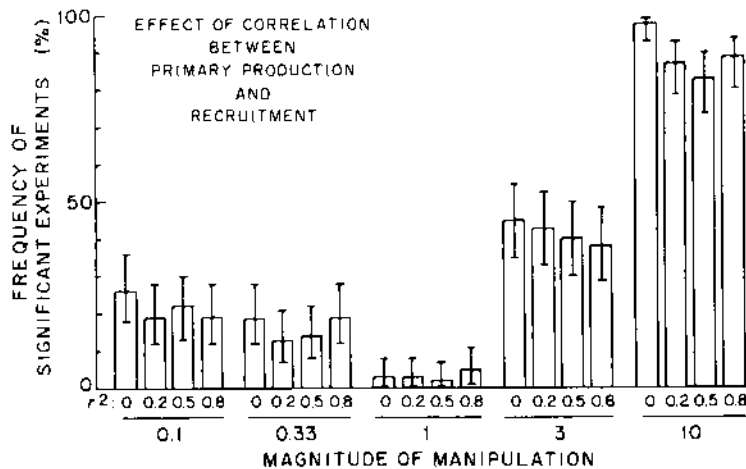


FIG. 6. Frequency of pulse experiments with significant *t* statistics ( $P < .05$ ) vs. magnitude of piscivore manipulation at four values of the  $r^2$  between primary production and fish recruitment. Data are means and 95% confidence intervals. All experiments used 10 lakes.

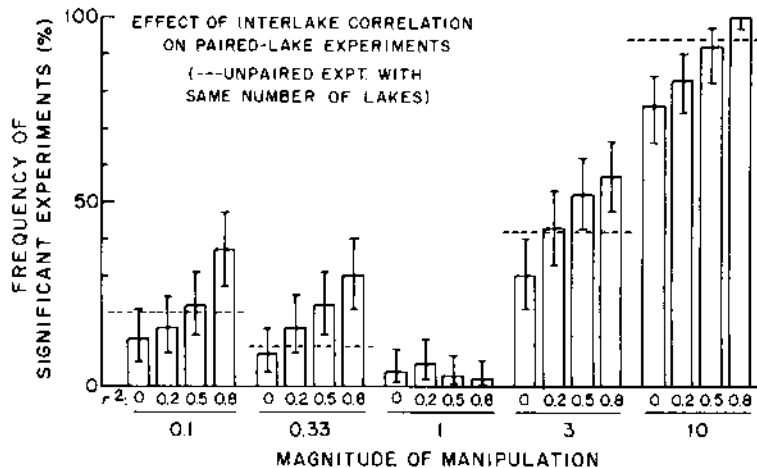


FIG. 7. Frequency of pulse experiments with significant  $t$  statistics ( $P < .05$ ) vs. magnitude of piscivore manipulation at four values of the  $r^2$  between interlake weather effects. Histogram shows means and 95% confidence intervals for experiments with five pairs of correlated lakes. Dashed lines show frequency of significant  $t$  statistics for experiments with 10 unpaired lakes and uncorrelated weather effects.

are consistent with model results showing that substantial manipulations are necessary. At the ecosystem level, the most convincing demonstrations of fish effects on phytoplankton have come from a whole-lake fish removal (Henrikson et al. 1980), a massive stocking of exotic piscivores causing 10- to 15-fold reductions in planktivory (Kitchell and Crowder 1986, Scavia et al. 1986), and addition of piscivores at a density near carrying capacity to a formerly piscivore-free system (Carpenter et al. 1987, Carpenter and Kitchell 1988). Less extensive manipulations have had lesser effects. For example,  $\approx 0.3\times$  planktivore removal caused an approximately twofold reduction in chlorophyll concentration that, while variable, appeared different from the premanipulation levels (Shapiro and Wright 1984). Piscivore reductions of  $\approx 0.5\times$  caused no detectable changes in chlorophyll concentrations

(McQueen et al. 1989). It is difficult to assess the magnitude of the piscivore addition by Benndorf et al. (1984), but zooplankton biomass increased about twofold; algal species composition changed, but there was no detectable change in total algal biomass. The magnitudes of the fish population changes in the lakes studied by Walters et al. (1987) were not quantified, but no effects on zooplankton or phytoplankton biomass were detected.

Because this model provides reasonable estimates of interannual variance in primary production, results are pertinent to any lake ecosystem experiment in which primary production is the response variate. Similar analyses are needed for other response variables, especially those at the community or population levels. Schindler (1987) has argued that species' population dynamics may be better indicators of ecosystem stress

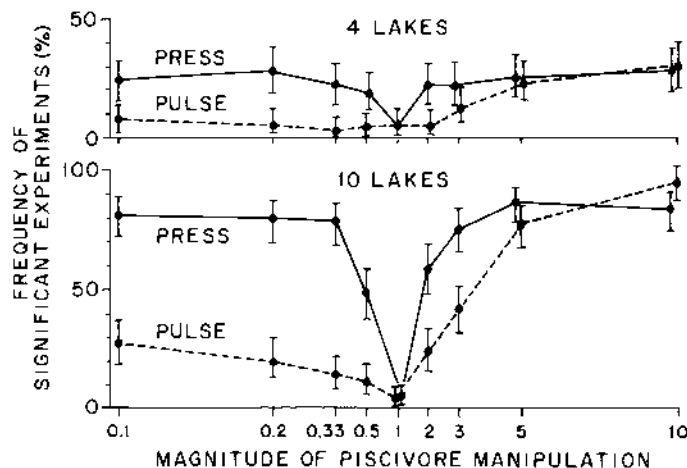


FIG. 8. Frequency of experiments with significant  $t$  statistics vs. magnitude of piscivore manipulation for press and pulse experiments with 4 and 10 lakes. Data are means and 95% confidence intervals.

than are traditional ecosystem processes such as primary production and nutrient cycling. While species responses may be relatively sensitive to ecosystem stress, they may also have greater variability than ecosystem processes. The trade-offs between sensitivity and variability of population, community, and ecosystem parameters are a critical issue in applied ecology and a fascinating challenge to basic ecologists. More complex food web models suggest that problems of topological indeterminacy will be a significant complication in analyses of community and population responses (Yodzis 1988).

#### *Pulse and press experiments*

Model results show that pulsed enhancements and removals of piscivores are not symmetric experimental designs. The asymmetry is an example of the hysteresis effects discussed by Carpenter et al. (1985). Piscivore additions have relatively long-lived effects, with duration proportional to the lifespan of the predator, a phenomenon termed "predatory inertia" (Stewart et al. 1981, Carpenter et al. 1985). Because the response persists and can be sampled for several years, it is more likely to be detected against the background variability of the ecosystem. In contrast, removal of <100% of the piscivores from a lake may be followed by relatively high recruitment as density-dependent limitations are relaxed (Carpenter et al. 1987). The rate of trophic ontogeny, or growth of the fishes to predominantly piscivorous size, then determines the duration of the ecosystem response (Werner and Gilliam 1984, Carpenter et al. 1985). Because the prepiscivorous phase is much shorter than the total lifespan of the fishes, the effects of piscivore removal will be of much shorter duration than those of a piscivore addition. Thus the experimenter has a relatively short time to sample the effects of a piscivore removal, and is less likely to detect the effects against the background variability of the ecosystem. Note that this argument does not apply to a total piscivore removal, in which postmanipulation recruitment would not occur.

Published studies, while sparse, appear to support the theoretical inference that effects of partial piscivore removals are harder to detect than those of the reciprocal piscivore enhancements. A piscivore removal of  $\approx 0.1 \times$  produced substantial but transient responses in phytoplankton composition, biovolume, and productivity that lasted less than a year (Peter Lake; Carpenter et al. 1987). A piscivore removal of  $\approx 0.5 \times$ , by winterkill, produced no detectable response in the phytoplankton (McQueen et al. 1988). Both of these studies showed lesser effects on the algae than both the piscivore enhancements in Lake Michigan (Kitchell and Crowder 1986, Scavia et al. 1986) and Tuesday Lake (Carpenter et al. 1987) and the fish removal from Lillestockelidsvatten (Henrikson et al. 1980).

Press enhancements and removals of piscivores were relatively symmetric experiments; i.e., any given en-

hancement experiment was as likely to succeed as its reciprocal removal experiment. In the press experiments, the algal response could be measured for a relatively long period of time and was more likely to be detected against the background variability of the ecosystem. The long duration of effects in press experiments overrides the difference between piscivore additions and removals that is evident in pulse experiments.

#### *Paired-lake experiments*

Regional correlations among lakes can substantially increase the sensitivity of ecosystem experiments. The value of paired-ecosystem experiments demonstrated by these simulations supports the intuitive notion of many experimenters who have used paired treatment and reference ecosystems (Likens 1985). Schindler et al. (1985) used two to eight reference lakes within the Experimental Lakes Area to assess the responses of an acidified lake. They used graphical comparisons rather than direct statistical tests. The data are convincing because of relatively small confidence intervals and similar temporal trends of the lakes, illustrating the experimental advantages that result from regional interlake correlation.

Paired designs have ecological advantages in addition to the statistical ones addressed by this paper. Closely associated lakes are likely to share many species in common, facilitating measurements of community change following manipulation. By pairing lakes over latitudinal or trophic gradients, the interactions of these gradients with food web effects can be assessed. Since the magnitude of piscivore effects across trophic and latitudinal gradients is currently controversial (McQueen et al. 1986, Crowder et al. 1988), experiments exploring the effects of these gradients would be especially valuable.

In the case of experimental designs that do not employ regionally correlated ecosystems (Figs. 5 and 6), my analyses probably underestimate the numbers of lakes needed to achieve significant results. At least two factors contribute to the underestimation. First, unless very similar lakes are used, interlake variance in primary production will exceed interannual variance within any given lake, because of interlake differences in nutrient loading (Schindler et al. 1978) and morphometry (Carpenter 1983). For example, the cv of annual primary production for the reference lakes used by Schindler et al. (1985) was 37%, much greater than the interannual cv of 20% used in the simulations reported here. However, correlated responses to regional weather patterns mitigate the effects of interlake variance. The paired-lake design is one effective way for experimenters to take advantage of interlake correlation. The second factor that entails greater sample sizes than those estimated here is the need to estimate time-treatment interactions using "staircase" experimental designs (Walters et al. 1988). My point about the need

for substantial and sustained manipulations applies with equal force to staircase experiments.

#### *Implications for ecosystem experiments*

The main conclusion is that relatively strong and sustained manipulations must be employed to reliably detect ecosystem responses. In the case of food web manipulations where primary productivity is the dependent variate,  $\approx 10$  replicate lakes are needed, and manipulations should enhance piscivore biomass at least  $5\text{--}10\times$  for several years to have a reasonably high probability of detecting responses in primary production. Similar numbers of replicates and comparably strong manipulations will be needed for whole-lake experiments on contaminants or nutrients, because these studies must also overcome the natural variability of lake primary production (Carpenter and Kitchell 1987, Schindler 1987). In fact, successful nutrient enrichment experiments have increased nutrient loads to manipulated lakes by 8 to  $15\times$  and sustained them for many years (Schindler et al. 1978). Whole-lake acidification experiments have increased hydrogen ion concentrations  $50\text{--}100\times$  for periods of many years (Schindler et al. 1985, Watras and Frost 1988). These manipulations are no larger than those commonly used in laboratory experiments, relative to the known variance of the experimental systems. Effective experimentation on complex and variable systems requires substantial and sustained manipulations.

The possibility of type II error (i.e., accepting the hypothesis that the manipulation had no effect, when in fact that hypothesis is false) is a critical one for all ecosystem experiments. All ecosystem experimenters face a shortage of suitable replicate systems and a shortage of sufficient and sustained funding for these relatively expensive and long-term studies. Therefore, it is tempting to use minimal replication (e.g., duplicates); consequently the probability of an erroneous nonsignificant result is very large. Such errors may have enormous costs in wasted research effort and faulty advice to resource managers.

In many cases, replicated whole-ecosystem experiments will simply be impossible. For example, experimental lakes are not common. Two prominent experimental reserves, Canada's Experimental Lakes Area (ELA) and the University of Notre Dame Environmental Research Center (UNDERC), contain only 46 and 27 lakes, respectively. Only certain subsets of these lakes constitute reasonable groups of replicates. In view of the large number of important experiments that remain to be done, and the need to sustain these experiments for many years, there are not enough lakes to replicate adequately.

Since sufficient replicate ecosystems will usually be unavailable, substantial and sustained manipulations in unreplicated paired-ecosystem experiments will continue to be the best available approach. Time series from these unreplicated experiments can be tested sta-

tistically for changes after manipulation, although the tests alone cannot prove that manipulation caused the changes (Box and Tiao 1975, Box et al. 1978, Stewart-Oaten et al. 1986, Frost et al. 1988, Carpenter et al. 1989). Performance of such experiments in different systems arrayed across major ecological gradients would be equivalent to carrying out large-scale paired-system experiments staggered in time and performed collectively by many scientists. Such collections of unreplicated experiments may be the most cost-effective, reliable, and rapid means of experimentation at the ecosystem level. There is great need for statistical guidance and innovations to improve the design and interpretation of large-scale ecological experiments.

#### ACKNOWLEDGMENTS

This work benefitted from discussions at the Notre Dame Workshop on Complex Interactions in Lake Communities and at the Second Cary Conference on Ecosystem Studies. A number of colleagues, who may or may not agree with my conclusions, have influenced my thoughts on these issues, especially L. Crowder, R. Drenner, T. Frost, S. Hurlbert, J. Kitchell, and D. McQueen. I thank T. Frost, J. Kitchell, D. McQueen, D. Schindler, G. Sprules, C. Walters, and two anonymous reviewers for comments on the manuscript. This work was supported by NSF grant BSR-86-06271 and the W. V. Kaeser Visiting Scholar Fund, University of Wisconsin-Madison.

#### LITERATURE CITED

- Bender, E. A., T. J. Case, and M. E. Gilpin. 1984. Perturbation experiments in community ecology: theory and practice. *Ecology* 65:1-13.
- Benndorf, J., H. Kneschke, K. Kossatz, and E. Penz. 1984. Manipulation of the pelagic food web by stocking with predaceous fishes. *Internationale Revue der gesamten Hydrobiologie* 69:407-428.
- Box, G. E. P., W. G. Hunter, and J. S. Hunter. 1978. *Statistics for experimenters*. John Wiley and Sons, New York, USA.
- Box, G. E. P., and G. C. Tiao. 1975. Intervention analysis with applications to economic and environmental problems. *Journal of the American Statistical Association* 70:70-79.
- Carpenter, S. R. 1983. Lake geometry: implications for sediment accretion rates. *Journal of Theoretical Biology* 105:273-286.
- Carpenter, S. R., editor. 1988a. *Complex interactions in lake communities*. Springer-Verlag, New York, New York, USA.
- Carpenter, S. R. 1988b. Transmission of variance through lake food webs. Pages 119-138 in S. R. Carpenter, editor. *Complex interactions in lake communities*. Springer-Verlag, New York, New York, USA.
- Carpenter, S. R., T. M. Frost, D. Heisey, and T. K. Kratz. 1989. Randomized intervention analysis and the interpretation of whole-ecosystem experiments. *Ecology* 70, *in press*.
- Carpenter, S. R., and J. F. Kitchell. 1987. The temporal scale of variance of limnetic primary production. *American Naturalist* 129:417-433.
- Carpenter, S. R., and J. F. Kitchell. 1988. Strong manipulations and complex interactions: consumer control of lake productivity. *BioScience* 38:764-769.
- Carpenter, S. R., J. F. Kitchell, and J. R. Hodgson. 1985. Cascading trophic interactions and lake productivity. *BioScience* 35:634-639.

- Carpenter, S. R., J. F. Kitchell, J. R. Hodgson, P. A. Cochran, J. J. Elser, M. M. Elser, D. M. Lodge, D. Kretchmer, X. He, and C. N. von Ende. 1987. Regulation of lake primary productivity by food web structure. *Ecology* **68**:1863-1876.
- Carpenter, S. R., P. R. Leavitt, J. J. Elser, and M. M. Elser. 1988. Chlorophyll budgets: effect of food web manipulation. *Biogeochemistry* **6**:79-90.
- Crowder, L. B., R. W. Drenner, W. C. Kerfoot, D. J. McQueen, E. L. Mills, U. Sommer, C. N. Spencer, and M. J. Vanni. 1988. Food web interactions in lakes. Pages 141-160 in S. R. Carpenter, editor. *Complex interactions in lake communities*. Springer-Verlag, New York, New York, USA.
- Frost, T. M., D. L. DeAngelis, T. F. H. Allen, S. M. Bartell, D. J. Hall, and S. Hurlbert. 1988. Scale in the design and interpretation of aquatic community research. Pages 229-260 in S. R. Carpenter, editor. *Complex interactions in lake communities*. Springer-Verlag, New York, New York, USA.
- Goldman, C. R., and A. de Amezaga. 1984. Primary productivity and precipitation at Castle Lake and Lake Tahoe during 24 years. *Internationale Vereinigung für theoretische und angewandte Limnologie, Verhandlungen* **22**:591-599.
- Henrikson, L., H. G. Nyman, H. G. Oscarson, and J. A. E. Stenson. 1980. Trophic changes, without changes in the external nutrient loading. *Hydrobiologia* **68**:257-263.
- Hurlbert, S. H. 1984. Pseudoreplication and the design of ecological field experiments. *Ecological Monographs* **54**:187-211.
- Kerfoot, W. C., and A. Sih. 1987. Predation: direct and indirect impacts on aquatic communities. University Press of New England, Hanover, New Hampshire, USA.
- Kitchell, J. F., S. M. Bartell, S. R. Carpenter, D. J. Hall, D. J. McQueen, W. E. Neill, D. Scavia, and E. E. Werner. 1988. Epistemology, experiments, and pragmatism. Pages 263-280 in S. R. Carpenter, editor. *Complex interactions in lake communities*. Springer-Verlag, New York, New York, USA.
- Kitchell, J. F., and L. B. Crowder. 1986. Predator-prey systems in Lake Michigan: model predictions and recent dynamics. *Environmental Biology of Fishes* **16**:205-211.
- Likens, G. E. 1985. An experimental approach for the study of ecosystems. *Journal of Ecology* **73**:381-396.
- Likens, G. E., F. H. Bormann, N. M. Johnson, D. W. Fisher, and R. S. Pierce. 1970. Effects of forest cutting and herbicide treatment on nutrient budgets in the Hubbard Brook watershed ecosystem. *Ecological Monographs* **40**:23-47.
- McQueen, D. J., M. R. S. Johannes, J. R. Post, T. J. Stewart, and D. R. S. Lean. 1989. Bottom-up and top-down impacts on freshwater pelagic community structure. *Ecological Monographs* **59**, *in press*.
- McQueen, D. J., J. R. Post, and E. L. Mills. 1986. Trophic relationships in freshwater pelagic ecosystems. *Canadian Journal of Fisheries and Aquatic Sciences* **43**:1571-1581.
- National Research Council. 1986. Global change in the geosphere-biosphere. National Academy Press, Washington, D.C., USA.
- Northcote, T. G. 1988. Fish in the structure and function of freshwater ecosystems: a "top-down" view. *Canadian Journal of Fisheries and Aquatic Sciences* **45**:361-379.
- Paine, R. T. 1980. Food webs: linkage, interaction strength and community infrastructure. *Journal of Animal Ecology* **49**:667-685.
- Rohlf, F. J., and R. R. Sokal. 1981. *Statistical tables*. W. H. Freeman, San Francisco, California, USA.
- Scavia, D., G. L. Fahnenstiel, M. S. Evans, D. Jude, and J. T. Lehman. 1986. Influence of salmonid predation and weather on long-term water quality trends in Lake Michigan. *Canadian Journal of Fisheries and Aquatic Sciences* **43**:435-443.
- Schindler, D. W. 1977. Evolution of phosphorus limitation in lakes. *Science* **195**:260-262.
- . 1987. Detecting ecosystem response to anthropogenic stress. *Canadian Journal of Fisheries and Aquatic Sciences* **44** (Supplement):6-25.
- . 1988. Experimental studies of chemical stressors on whole lake ecosystems. *Internationale Vereinigung für theoretische und angewandte Limnologie, Verhandlungen* **23**:11-41.
- Schindler, D. W., F. A. J. Armstrong, S. K. Holmgren, and G. J. Brunskill. 1971. Eutrophication of Lake 227, Experimental Lakes Area, Northwest Ontario, by addition of phosphate and nitrate. *Journal of the Fisheries Research Board of Canada* **28**:1763-1782.
- Schindler, D. W., E. J. Fee, and T. Ruzsyczynski. 1978. Phosphorus input and its consequences for phytoplankton standing crop and production in the Experimental Lakes Area and in similar lakes. *Journal of the Fisheries Research Board of Canada* **35**:190-196.
- Schindler, D. W., K. H. Mills, D. F. Malley, D. L. Findlay, J. A. Shearer, I. J. Davies, M. A. Turner, G. A. Linsey, and D. R. Cruikshank. 1985. Long-term ecosystem stress: the effects of years of acidification on a small lake. *Science* **228**:1395-1401.
- Shapiro, J., and D. I. Wright. 1984. Lake restoration by biomanipulation. *Freshwater Biology* **14**:371-383.
- Sokal, R. R., and F. J. Rohlf. 1981. *Biometry*. W. H. Freeman, San Francisco, California, USA.
- Stewart, D. J., J. F. Kitchell, and L. B. Crowder. 1981. Forage fishes and their salmonid predators in Lake Michigan. *Transactions of the American Fisheries Society* **110**:751-763.
- Stewart-Oaten, A., W. Murdoch, and K. Parker. 1986. Environmental impact assessment "pseudoreplication" in time? *Ecology* **67**:929-940.
- Walters, C. J., J. S. Collie, and T. Webb. 1988. Experimental designs for estimating transient responses to management disturbances. *Canadian Journal of Fisheries and Aquatic Sciences* **45**:530-538.
- Walters, C. J., E. Krause, W. E. Neill, and T. G. Northcote. 1987. Equilibrium models for seasonal dynamics of plankton biomass in four oligotrophic lakes. *Canadian Journal of Fisheries and Aquatic Sciences* **44**:1002-1017.
- Watras, C., and T. M. Frost. 1988. Little Rock Lake: perspectives on an experimental ecosystem approach to seepage lake acidification. *Archives of Environmental Contamination and Toxicology*, *in press*.
- Werner, E. E., and J. F. Gilliam. 1984. The ontogenetic niche and species interactions in size-structured populations. *Annual Review of Ecology and Systematics* **15**:393-425.
- Yodzis, P. 1988. The indeterminacy of ecological interactions as perceived through perturbation experiments. *Ecology* **69**:508-515.