

Fitting Predator-Prey Models to Time Series with Observation Errors

Author(s): S. R. Carpenter, K. L. Cottingham and C. A. Stow

Source: *Ecology*, Vol. 75, No. 5 (Jul., 1994), pp. 1254-1264

Published by: Wiley

Stable URL: <http://www.jstor.org/stable/1937451>

Accessed: 14-01-2017 15:45 UTC

## REFERENCES

Linked references are available on JSTOR for this article:

[http://www.jstor.org/stable/1937451?seq=1&cid=pdf-reference#references\\_tab\\_contents](http://www.jstor.org/stable/1937451?seq=1&cid=pdf-reference#references_tab_contents)

You may need to log in to JSTOR to access the linked references.

---

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact [support@jstor.org](mailto:support@jstor.org).

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at

<http://about.jstor.org/terms>



Wiley is collaborating with JSTOR to digitize, preserve and extend access to *Ecology*

## FITTING PREDATOR–PREY MODELS TO TIME SERIES WITH OBSERVATION ERRORS<sup>1</sup>

S. R. CARPENTER, K. L. COTTINGHAM, AND C. A. STOW  
*Center for Limnology, University of Wisconsin, Madison, Wisconsin 53706 USA*

**Abstract.** Fits of nonlinear difference equations to time series with observation errors were examined by stochastic simulation and analysis of plankton time series from two lakes. Even modest observation errors (e.g., coefficient of variation among replicate samples  $\approx 10\%$ ) cause errors in model identification and bias in parameter estimates. The latter problem can be corrected by estimation techniques that account for observation error, but model identification is difficult unless the state variables are manipulated. Without manipulation, statistical criteria tend to favor linear models, even when data are simulated by nonlinear processes. Methods that account for observation error produced satisfactory fits to time series of edible algae from two lakes over 7 yr. In Paul Lake, which has not been manipulated, the best fitting model included linear growth, a linear functional response for grazing loss, and an autoregressive moving average model for the errors. In manipulated Tuesday Lake, the best fitting model included linear growth and a nonlinear functional response. Experimental manipulations, or other substantial perturbations, may be essential for detection of nonlinearities in ecological interactions.

**Key words:** *difference equations; ecosystem experiments; error in variables; lakes; plankton; predator–prey models; time series.*

### INTRODUCTION

Forecasts based on ecological time series have many applications in ecology and resource management. Models fit to time series can be used to anticipate changes in water quality or resource stocks and evaluate potential effects of management actions (Jassby and Powell 1990, Hilborn and Walters 1992). Time series modeling may also be used to infer causes of temporal variability (Jassby et al. 1990), test consequences of natural or deliberate perturbations (Carpenter and Kitchell 1993, Rudstam et al. 1993), and compare different mechanisms for ecological interactions (Blau and Neely 1975, Morris 1990, Carpenter et al. 1994). Nevertheless, examples of model discrimination and forecasting using ecological time series are relatively few. With increased attention to the collection and preservation of long-term data sets, ecologists should have expanding opportunities to test hypotheses and develop predictive capabilities using time series (Likens 1989, Magnuson 1990, Risser 1991).

The problem of model fitting when both predictor and response variables are subject to error commonly arises in analyses of ecological models. Parameter estimation under “errors-in-variables” has been well explored in the statistical literature (Sprent 1990). In the simple linear regression model, failure to accommodate observation error on the predictor variable results in least squares and maximum likelihood slope estimators that are biased toward zero (Fuller 1987). Bi-

ased estimators also arise in multiple regression and nonlinear models, though the direction of bias is case specific (Seber and Wild 1989).

If the maximum likelihood expression is expanded to incorporate stochastic predictor observation error, then information regarding the variance of the observation error must be provided for optimal parameter estimates to be uniquely identified (Judge et al. 1985). Even if the variance of the predictor observation error is known, obtaining parameter estimates can still be difficult because there exists an unknown or incidental parameter associated with each observation of each predictor. The result is an optimization problem that increases in dimension with the amount of data. Several methodologies have been proposed to alleviate this problem by eliminating the incidental parameters. These include modified likelihood estimators that incorporate model sensitivity to the uncertain predictors (Clutton-Brock 1967, Chandler 1972), an analogous Bayesian approach in which the incidental parameters are integrated out of the joint posterior density function (Reilly and Patino-Leal 1981), and the use of James–Stein estimators in place of the erroneously observed predictors (Whittemore 1989).

In an ecological context, parameter bias resulting from a failure to accommodate observation error may encourage overexploitation of fish stocks (Walters 1986) and result in errors in choosing the most appropriate model from a series of alternatives (Hilborn and Walters 1992). Most ecological work on this problem has focused on situations in which the observation error variance is poorly known (Ludwig and Walters 1989). We address a situation in which observation error vari-

<sup>1</sup> Manuscript received 1 June 1993; revised 28 September 1993; accepted 10 October 1993.

ance is well estimated through replicate measurements, and the methodologies of Clutton-Brock (1967), Chandler (1972), and Reilly and Patino-Leal (1981) can be employed. We apply the method to a problem of current interest, the modeling of the functional response in predator-prey interactions. Using stochastic simulation, we show that alternative models can sometimes be distinguished if the state variables are manipulated experimentally. We compare three alternative models for phytoplankton-zooplankton interactions in two lakes.

### Functional response

The current debate about the functional response in predator-prey interactions is an example of an issue that could be illuminated by fitting models to ecological time series (Arditi and Ginzburg 1989, Hanski 1991, Oksanen et al. 1992). Discussion involves the general model for dynamics of prey ( $V$ ) and predator ( $P$ ):

$$V_{t+1} = R(V_t) - f(V_t, P_t) \quad (1)$$

$$P_{t+1} = af(V_t, P_t) - mP_t, \quad (2)$$

where  $R$  is the recruitment function for the prey,  $f$  is the predation function,  $a$  is a conversion efficiency, and  $m$  is the predator's mortality rate. Alternative predation functions have very different implications for responses of ecosystems to enrichment and the temporal variability of long-term dynamics (Berryman 1992).

In prey-dependent models, the Type 2 (or asymptotic) predation function is

$$f(V_t, P_t) = cV_tP_t/(h + V_t) \quad (3)$$

(Holling 1959). Here  $c$  is the maximum consumption rate and  $h$  is the half-saturation coefficient (the prey density where half the maximum consumption rate occurs). If the prey recruitment function is density dependent, the attractor for the continuous-time version of this model is a stable point for systems with low productivity or a limit cycle for systems with high productivity (Rosenzweig 1971). At equilibrium, enrichment increases the predator biomass but not the prey biomass (Arditi and Ginzburg 1989).

In ratio-dependent models, the asymptotic predation function is

$$f(V_t, P_t) = c(V_t/P_t)P_t/[h + (V_t/P_t)] \quad (4a)$$

or equivalently

$$f(V_t, P_t) = cV_tP_t/[(hP_t) + V_t] \quad (4b)$$

(Berryman 1992). The parameter  $c$  has the same meaning as in the prey-dependent equation, while  $h$  is the prey-predator ratio at which half the maximum consumption rate occurs. In the continuous-time version of this model, the attractor is less likely to be destabilized by enrichment (Berryman 1992). At equilibrium, enrichment increases both the predator and the

prey biomasses (Arditi and Ginzburg 1989, Arditi et al. 1991).

The contrasting stability properties and responses to enrichment of these models have significant implications for ecological concepts and ecosystem management (Hanski 1991, Slobodkin 1992). Tests of the models are thus of great interest. Some researchers have turned to plankton for relevant data, because phytoplankton-zooplankton dynamics are rapid and experiments are possible at a variety of scales. Laboratory experiments offer some evidence for both models (Porter et al. 1982, Arditi and Saiah 1992), but it is not certain if the small-scale results apply at the ecosystem scale (Frost et al. 1988).

At the whole-lake scale, comparative studies appear to support the ratio-dependent model by showing positive correlations between algal biomass and zooplankton biomass across gradients of productivity in lakes (Ginzburg and Akcakaya 1992). However, these correlations can also be explained by: (1) increasing dominance of inedible algae in more productive lakes (McCauley et al. 1988, Watson and McCauley 1988); (2) increasing zooplankton death rates in more productive lakes (Leibold 1990, Gatto 1991, Sarnelle 1992); or (3) algae-zooplankton dynamics that are not at equilibrium (Carpenter 1988, Carpenter and Kitchell 1993). Other ecosystem comparisons indicate that trends in zooplankton size affect the relationship between algal biomass and zooplankton biomass (Pace 1984, Carpenter et al. 1991).

These difficulties may be avoided by fitting alternative models to time series of edible algae from lakes. This approach does not assume equilibrium conditions. By fitting models to the algal time series, we eliminate complications from shifts in zooplankton death rates and zooplankton age structure. Possible effects of inedible algae are avoided by measuring time series for small algae that are eaten by a broad spectrum of herbivores. However, the need to consider both observation error and process error could make the fitting and comparison of these nonlinear models very difficult. This paper explores approaches to that problem.

### METHODS

We analyze both stochastic simulations and time series from an ecosystem experiment. In the stochastic simulations, time series of predator and prey were generated using one of two expressions for the predation function: prey dependent (Eq. 3) or ratio dependent (Eq. 4b). Both process and observation errors were included in the output time series. Simulated data were used for two purposes. First, we compared alternative estimators for model parameters. Second, we assessed our ability to identify the correct model from a set of alternatives.

In the whole-lake experiment, alternative models were fit to time series from a manipulated lake and a reference lake. To compare the models, we examined

the residuals and prediction errors obtained by bootstrapping (Hilborn and Walters 1992, Leger et al. 1992).

### Stochastic simulations

Time series were simulated for the model

$$V_{t+1} = \{rV_t - sV_t^2 - f[V_t, P_t]\}\exp[v_{Vt}] \quad (5)$$

$$P_{t+1} = \{af[V_t, P_t] - mP_t\}\exp[v_{Pt}] \quad (6)$$

$$V_{\text{obs},t} = V_t \exp[\omega_{Vt}] \quad (7)$$

$$P_{\text{obs},t} = P_t \exp[\omega_{Pt}]. \quad (8)$$

In these equations,  $r$  is the prey growth rate (corresponding to  $r + 1$  in the discretized version of the logistic differential equation; May 1981);  $s$  is self-limitation of prey;  $a$  is the conversion efficiency;  $m$  is the mortality rate of the predator;  $v$  are the process errors; and  $\omega$  are the observation errors.  $V_{\text{obs}}$  and  $P_{\text{obs}}$  are the observed values of the time series, influenced by both process and observation errors. Process errors represent effects not included in the model that cause the dynamics to depart from model predictions. Observation errors represent variability in the data arising from sampling and analysis.

Time series were generated using both prey-dependent and ratio-dependent predation functions. Parameter values used in the simulations appear in Table 1. Values chosen were typical of zooplankton-phytoplankton systems (Carpenter 1992). Initial conditions for the simulations were near equilibrium. Except where stated otherwise, time series of 50 observations were simulated. This length was chosen to mimic a typical plankton study spanning  $\approx 1$ –4 yr of weekly to monthly observations.

Errors of  $V_{\text{obs}}$  and  $P_{\text{obs}}$  were distributed lognormally, a common property of ecological data. Deterministic components were multiplied by the exponential of errors derived from a normal pseudorandom generator with zero mean and standard deviations shown in Table 1. It is convenient to think of the distributions of  $V_{\text{obs}}$  and  $P_{\text{obs}}$  in terms of their coefficients of variation (cv = standard deviation divided by the mean, calculated for untransformed data) because the cv is a constant for lognormally distributed variates. The mean ( $m_z$ ) and standard deviation ( $s_z$ ) of a lognormal variate  $Z = \exp(X)$  where  $X$  is a normal variate with mean  $m_x$  and standard deviation  $s_x$  are

$$m_z = \exp(m_x)\exp(s_x^2/2)$$

$$s_z = m_z[\exp(s_x^2) - 1]^{0.5}.$$

The cv of untransformed data is simply  $s_z/m_z = [\exp(s_x^2) - 1]^{0.5}$ . For the standard deviations given in Table 1, the cv for observation errors of untransformed data is  $\approx 10\%$  for  $V$  and  $20\%$  for  $P$ .

Simulations with very low observation errors were performed to test our ability to discriminate alternative models with precise data. For these simulations, the

TABLE 1. Parameter values used in the simulations.

Parameter		Equation no.	Value
Name	Symbol*		
Conversion efficiency	$a^P$	2,6,5–8	0.6
	$a^R$	2,3,5–8	0.8
Max. consumption rate	$c^P$	3,4	5
	$c^R$	3,4	25
Half-saturation coefficient	$h^P$	3,4	600
	$h^R$	3,4	60
Predation mortality rate	$m$	2,6	0.001
Prey growth rate	$r$	5	1.8
Self-limitation of prey	$s$	5	0.0018
Process errors	$v_V, v_P$	5–6	
SD of $v_V$ and $v_P$			0.017
Observation errors	$\omega_V, \omega_P$	7–8	
SD of $\omega_V$			0.10
SD of $\omega_P$			0.20

\* P = prey-dependent. R = ratio-dependent.

standard deviations of  $\omega_V$  and  $\omega_P$  were set at 0.0001. In this case, the cv of the observation errors is  $\approx 0.01\%$ .

Simulations with perturbations of the state variables were performed to test our ability to discriminate alternative models in ecosystem experiments. In experimental simulations, prey biomass was increased to 400 during time steps 10–20 and predator biomass was increased to 100 (prey-dependent simulations) or 300 (ratio-dependent simulations) during steps 30–40.

### Model comparisons

Contrasting versions of Eq. 1 were fit to the time series  $V_{\text{obs}}$  and  $P_{\text{obs}}$  by methods that account for observation error (Clutton-Brock 1967, Chandler 1972, Reilly and Patino-Leal 1981). In the simulation studies, three versions of Eq. 5 were compared. These included (1) a prey-dependent linear predation term  $f = cV_tP_t$ , (2) the asymptotic prey-dependent predation term (Eq. 3), or (3) the asymptotic ratio-dependent predation term (Eq. 4b). For the fits to field data, we considered these three models plus three density-independent models formed by setting  $s$  of Eq. 5 to zero. In all cases, models were fit by minimizing a loss function by the simplex method (Nelder and Mead 1964) to a tolerance of  $10^{-4}$ .

In the fits to field data, residuals were sometimes autocorrelated. Autocorrelation in the residuals suggests time-series bias in the parameter estimates (Hilborn and Walters 1992). When autocorrelations or partial autocorrelations of residuals were significant ( $P < 0.05$ ), we added autoregressive moving average (ARMA) terms to the models (Appendix). ARMA parameters were fit jointly with other parameters of the model to remove this autocorrelation and improve parameter estimates. We used the minimum number of ARMA terms necessary to produce nonsignificant autocorrelation and partial autocorrelation functions for the residuals. At most, our data required one autoregressive parameter and one moving average parameter.

Models were compared using several criteria. Be-

cause the estimators require weighted loss functions where the weightings differ among models, likelihoods could not be compared directly (Edwards 1992).

In the simulation studies, we compared goodness of fit using the standard deviation of residuals. We also compared models using the sum of squared residuals weighted by the observation variances (i.e.,  $\sum g_t^2/[BQ_tB']$ , where  $g_t$  is the residual at time  $t$ ,  $Q_t$  is the observation covariance matrix at time  $t$ , and  $B$  is a unit vector) and the negative log likelihood of the residuals. These criteria usually identified the same model as best fitting, so here we present the simplest criterion, the residual standard deviation.

In the fits to field data, we compared the models on the basis of the residual standard deviation, the correlation between predictions and observations, and the standard deviation of predictions. The standard deviation of predictions was obtained by bootstrapping (Hilborn and Walters 1992, Leger et al. 1992), as follows. Pseudovalues were generated by adding randomly selected (with replacement) residuals to model estimates. Parameters were estimated for 100 sets of pseudovalues, and the covariance matrix  $C$  of the bootstrapped parameters was calculated. A predictor vector  $B$  was formed from the partial derivative of the model with respect to each parameter, evaluated at the mean value of each predictor and the point estimates of the parameter. The prediction standard deviation was then calculated as  $(BCB')^{0.5}$ . This quantity can be understood as the standard deviation of a prediction at the centroid of the data.

#### *Fitting procedure*

Several approaches have been used to fit dynamic models to time series when both process and observation errors must be considered. One approach is to fit the models as if all the errors were process errors, or as if all the errors were observation errors, and then assess the magnitude of the bias through stochastic simulation (Hilborn and Walters 1992). This method is advantageous when process and observation errors cannot be estimated separately. An alternative is to develop a loss function for the particular model and data set, and fit the model by minimizing the loss function (Ludwig and Walters 1989). This method requires an independent estimate of observation error, e.g., through replicate samples.

The literature offers three loss functions that could be minimized with respect to parameter choices to fit predator-prey models. We compared all three of these loss functions to ordinary least squares, in which the sum of squared residuals is minimized. In the equations below,  $V$  and  $P$  refer to the observed mean values of predator and prey.

The first and simplest expression is similar to weighted least squares, except that the weightings are determined by the variances of both predictor and response

variables, rather than just the variance of the response variable (Chandler 1967, Reilly and Patino-Leal 1981):

$$\sum g_t^2(B_t Q_t B_t')^{-1}. \quad (9)$$

The summation is over time points  $t$ .  $Q_t$  is the variance-covariance matrix for untransformed replicate observations of  $V$  and  $P$  at time  $t$ . The term  $g$  is the residual at each time step

$$g_t = V_{t+1} - rV_t - sV_t^2 - f(V_t, P_t) \quad (10)$$

and  $B_t$  is a vector of partial derivatives of  $g$  with respect to each of the three variates  $V_{t+1}$ ,  $V_t$ , and  $P_t$ .

The second loss function is more similar to a negative log likelihood function in that the scaling factor of the normal distribution has been retained (Clutton-Brock 1967):

$$\sum 0.5[g_t^2(B_t Q_t B_t')^{-1} + \log(2\pi B_t Q_t B_t')]. \quad (11)$$

Expressions 9 and 11 are both based on the normal distribution, and therefore may not be a good approximation for the errors in our lognormally distributed data. On the other hand, the fact that observation errors  $Q_t$  are estimated independently (by replicate sampling) at each time step allows the variances to increase as the mean increases. This is a key property of the lognormal distribution for ecological data, and explains why we obtained accurate parameter estimates from Eqs. 9 and 11 (see *Results*).

The third loss function is a negative log likelihood based on the lognormal distribution:

$$\sum [0.5(\log g_L)^2(B_L Q_L B_L')^{-1} + [0.5 \log(2\pi g_L^2 B_L Q_L B_L')]]. \quad (12)$$

The residuals  $g_L$  are now defined as

$$g_L = [rV_t - sV_t^2 - f(V_t, P_t)]/V_{t+1}. \quad (13)$$

$B_L$  is now the vector of partial derivatives of  $\log g_L$  with respect to each of the three means  $V_{t+1}$ ,  $V_t$ , and  $P_t$ .  $Q$  is the variance-covariance matrix for log transformed replicates, which is assumed constant over time. Although Expression 12 is the most appropriate probability model for our data, it is difficult to implement numerically because the numerator in Eq. 13 can be negative, causing errors when the program calls for the log of a negative number. This problem can be overcome by using constrained optimization techniques, direct search, or simulated annealing (Bohachevsky et al. 1986). However, these methods are relatively slow and were therefore disadvantageous for the simulation and bootstrapping studies we performed. For other applications, minimization of Eq. 12 by direct search or simulated annealing may be productive.

#### *Limnological data*

Biomasses of edible phytoplankton and cladoceran herbivores were estimated at weekly intervals during

summer stratification for 7 yr (1984–1990) in two lakes (Carpenter and Kitchell 1993). The food web of Tuesday Lake was manipulated twice during this period. In 1985, piscivorous largemouth bass were added and planktivorous minnows were removed. In autumn 1986, bass were removed. Minnows were reintroduced in the spring of 1987. These manipulations caused dramatic changes in the zooplankton. Herbivore biomass was low in 1984 and the first half of 1985 and in 1989–1990. Herbivore biomass was high from mid-1985 through most of 1988. Nutrient levels were not manipulated. With respect to the models studied in this paper, the Tuesday Lake experiment is a press manipulation of  $P$  (the zooplankton) effected by the fishes, while the nutrient supply to  $V$  (the phytoplankton) was not altered directly. Throughout 1984–1990, Paul Lake was left undisturbed as a reference system.

Phytoplankton samples were collected from three depths using a Van Dorn bottle and pooled for enumeration (Carpenter and Kitchell 1993). Concentration and biovolume (i.e., the volume of protoplasm excluding the integument and surficial structures) were determined for each taxon (Carpenter and Kitchell 1993). Edible phytoplankton biomass was calculated as the combined biomass of all phytoplankters with biovolume less than that of a 30- $\mu\text{m}$  diameter sphere. These algae are generally edible by the grazer taxa included in our estimates of herbivore biomass (Bergquist and Carpenter 1986). Observation error cv was estimated as the cv among chlorophyll  $a$  concentrations measured in the same three samples that were pooled for phytoplankton enumeration (Carpenter and Kitchell 1993). We assumed that the chlorophyll cv was equal to the edible biovolume cv.

Zooplankton samples were collected by pooling two vertical hauls of a 80- $\mu\text{m}$  mesh net (Carpenter and Kitchell 1993). Animals were enumerated by species, and biomasses were calculated using length–mass regressions for each species (Carpenter and Kitchell 1993). Here we report biomass of cladoceran herbivores, the most important of which are *Bosmina longirostris*, *Daphnia parvula*, *D. pulex*, *D. rosea*, *Diaphanosoma birgei*, and *Holopedium gibberum*. Observation error cv was estimated as the cv among replicate vertical profiles measured during studies of net efficiency and diel vertical migration that could account for spatial variability in zooplankton densities (Carpenter and Kitchell 1993).

## RESULTS

### Comparison of estimators

We compared the loss functions (Expressions 9, 11, and 12) using simulated data. All three loss functions gave parameter estimates close to the actual values (Fig. 1). Ordinary least squares also gave good estimates of  $r$  and  $s$ , parameters in which the model was linear. However, ordinary least squares yielded biased

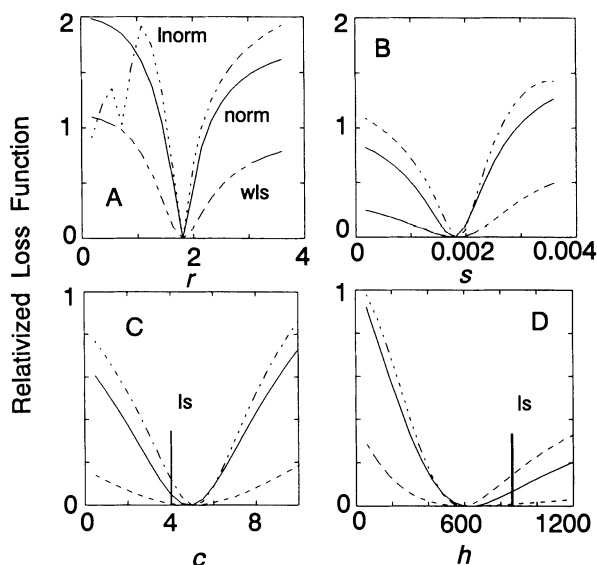


FIG. 1. Three loss functions vs. parameter values for time series of 100 values simulated with the Type 2 prey-dependent model and parameters of Table 1 (except that the standard deviation of  $\omega_p$  was 0.10). Ordinates were relativized so graphs could be overlaid.  $\cdots$ —, weighted least squares (wls, Eq. 9); —, normal likelihoods (norm, Eq. 11); - - -, lognormal likelihoods (lnorm, Eq. 12). For the nonlinear parameters  $c$  and  $h$ , the location of the least squares (ls) estimates is marked with a vertical line. Parameters: (A)  $r$ , actual value = 1.8; (B)  $s$ , actual value = 0.0018; (C)  $c$ , actual value = 5; and (D)  $h$ , actual value = 600.

estimates of parameters  $c$  and  $h$ , in which the model was nonlinear (Fig. 1C, D). Of the three apparently unbiased estimators, Expression 9 was fastest to compute, while Expression 11 tended to converge in the fewest iterations. Expression 12 was not well suited to simplex minimization due to numerous local minima and domain errors when  $g_L$  is negative.

Because Expressions 9, 11, and 12 yielded nearly identical and apparently unbiased parameter estimates for our models, we assumed that any of the three loss functions could be used for model comparisons. Expression 11 was minimized with respect to parameter choices for the following results.

### Model identification

Time series generated by the models show fluctuations similar to those of ecological data (Fig. 2). The experimental perturbations caused notable departures from the baseline fluctuations.

In simulations using the prey-dependent Type 2 (PD2) model and low observation error ( $cv = 0.01\%$ ), the correct model was reliably identified (Fig. 3A). The PD2 model had a much lower residual standard deviation ( $s_E$ ) than the ratio-dependent Type 2 (RD2) and prey-dependent Type 1 (PD1) alternatives (Fig. 3A). With observation error  $cv = 10\%$ , the  $s_E$  values were similar for all three models (Fig. 3B). The PD2 model fit slightly better than the RD2 model in  $>90\%$

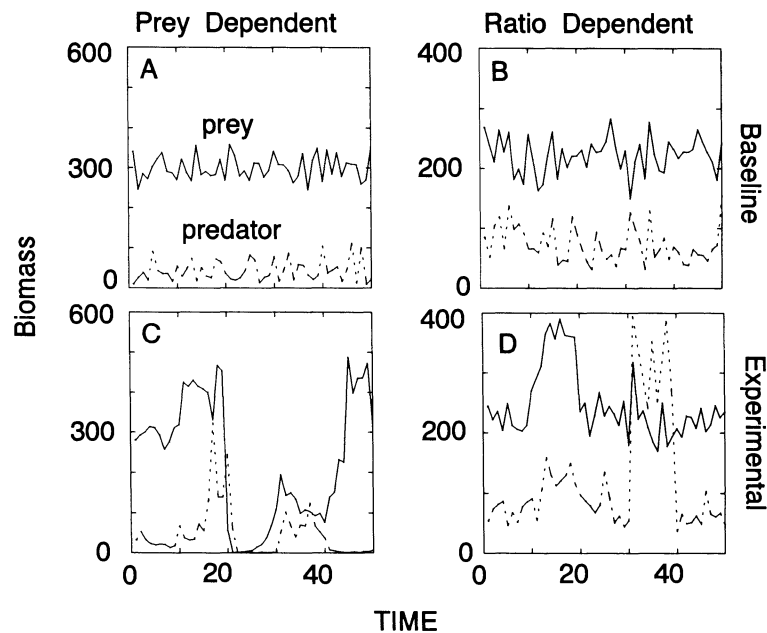


FIG. 2. Simulations of predator ( $P_{obs}$ ,  $\cdots$ ) and prey ( $V_{obs}$ ,  $\text{—}$ ) using the parameter values in Table 1. All models include logistic prey growth and a Type 2 functional response. (A) Prey-dependent predation under baseline conditions (i.e., no experimental perturbation of the predator population). (B) Ratio-dependent predation under baseline conditions. (C) Prey-dependent predation with prey biomass increased during time steps 10–20 and predator biomass increased during time steps 30–40. (D) Ratio-dependent predation with experimental manipulations as in panel (C).

of the simulations. However, the PD2 model did not fit as well as the PD1 model in >90% of the simulations. With observation error  $cv = 10\%$  and experimental perturbation of the state variables, the differences in  $s_E$  were larger (Fig. 3C). The PD2 model fit better than the RD2 model in >80% of the simulations and always fit better than the PD1 model.

When data were simulated with the RD2 model with low observation error, the correct model was identified in some cases (Fig. 3D). The RD2 model fit better than the PD2 model in >90% of the simulations (Fig. 3D). However, in most cases the PD1 model fit about as well as the RD2 model (Fig. 3D). With observation error  $cv = 10\%$ , the three models fit equally well (Fig. 3E). Although differences in  $s_E$  were small, the RD2 model fit better than the PD2 model in >80% of the cases, and fit better than the PD1 model in >60% of the cases. With observation error  $cv = 10\%$  and experimental perturbation of the state variables, comparisons with the PD2 model became bimodal (Fig. 3F). In about half the simulations the RD2 model was correctly identified. In the other half of the simulations, the RD2 and PD2 models fit equally well. The PD1 model fit about as well as the RD2 model in all simulations (Fig. 3F).

#### Fits to limnological data

Observation errors for the limnological data were similar to those used in the simulations. In Paul Lake, the standard deviations of  $\omega_V$  and  $\omega_P$  were 0.110 and

0.149, respectively, corresponding to cvs of  $\approx 11$  and 15%. In Tuesday Lake, the standard deviations of  $\omega_V$  and  $\omega_P$  were 0.080 and 0.168, respectively, corresponding to cvs of  $\approx 8$  and 17%. In fits of the models to the data from Paul and Tuesday Lakes, the parameters of the PD2 functional response had large variances and were nearly perfectly correlated. The residuals and residual variances of the PD2 and PD1 models were nearly identical. We concluded that the PD1 and PD2 models were redundant for these data and used the model with fewer parameters (PD1) in subsequent analyses.

In Paul Lake, models without density dependence fit better than models with density dependence (Table 2). In models with density dependence,  $r$  and  $s$  were highly correlated, and bootstrap estimates of the parameters often did not converge. Therefore, the prediction standard deviation could not be calculated. In the comparison of models without density dependence, the model with the PD1 functional response fit better on the basis of all criteria: lower residual  $s$ , higher correlation of predictions and observations, and lower prediction  $s$ . One-step-ahead predictions of this model were very close to observations (Fig. 4B).

Parameter estimates for the best fitting model in Paul Lake were prey growth rate  $r = 1.14 \text{ wk}^{-1}$  and maximum prey consumption rate  $c$ , per week per milligram of zooplankton per square metre,  $= 0.377 \text{ wk}^{-1} \cdot (\text{mg}/\text{m}^2)^{-1}$ . At typical biomasses (edible algae  $\approx 0.2 \text{ mg/L}$ , zooplankton  $\approx 0.5 \text{ mg}/\text{m}^2$ ), net weekly growth is  $\approx 0.23$

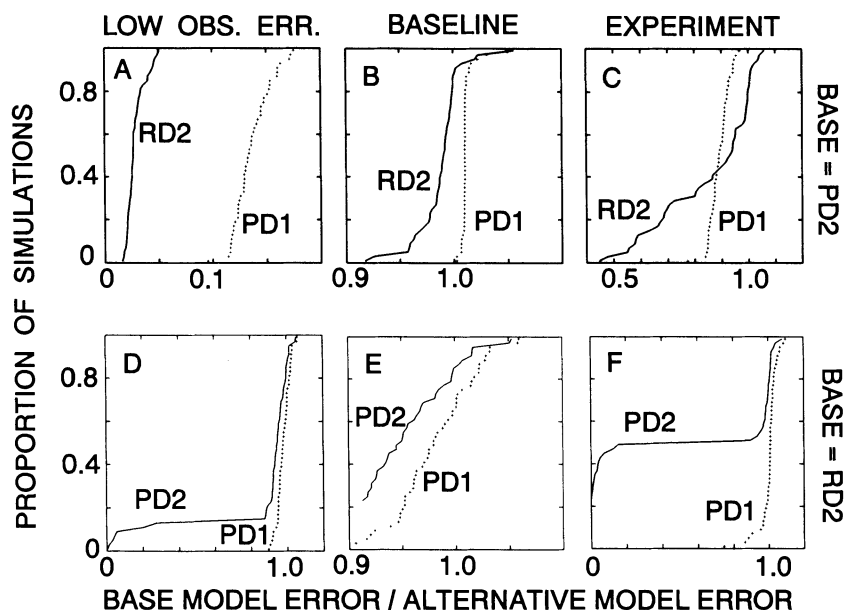


FIG. 3. Quantile plots of the base model error standard deviation divided by the alternative model standard deviation in fits to simulated data. Each line shows the proportion of simulations having a lower ratio of error standard deviations than the abscissa value. Ratios  $< 1$  favor the correct (base) model. (A) Prey-dependent Type 2 (PD2) base model vs. ratio-dependent Type 2 (RD2, —) and prey-dependent Type 1 (PD1, ·····) alternatives, observation error  $cv = 0.01\%$ . (B) Same comparison as panel (A) except observation error  $cv = 10\%$ . (C) Same conditions as panel (B) with experimental manipulations of state variables. (D) RD2 base model vs. PD2 (—) and PD1 (·····) alternatives, observation error  $cv = 0.01\%$ . (E) Same comparison as panel (D) except observation error  $cv = 10\%$ . (F) Same conditions as panel (E) with experimental manipulations of state variables.

mg/L. Weekly grazing loss is  $\approx 0.04$  mg/L, or  $\approx 17\%$  of growth.

In Tuesday Lake, the models with RD2 functional responses fit somewhat better than those with PD1 functional responses (Table 2). The model with RD2 functional response and no density dependence appeared to be best. This model had a higher residual  $s$  than the density-dependent RD2 model. However, it did not require extra parameters to correct autocorrelations in the errors and had a markedly lower prediction  $s$  than the alternatives. Predictions of the density-independent RD2 model had the highest correlation

with observations. One-step-ahead predictions of this model were very close to observations (Fig. 4D).

Parameter estimates for the best fitting model in Tuesday Lake were  $r = 1.11 \text{ wk}^{-1}$ ,  $c = 0.377 \text{ mg} \cdot \text{L}^{-1} \cdot \text{wk}^{-1} \cdot (\text{mg}/\text{m}^2)^{-1}$ , and  $h$ , in milligrams of algae per litre per milligram of zooplankton per square metre,  $= 0.562 \text{ mg} \cdot \text{L}^{-1} \cdot (\text{mg}/\text{m}^2)^{-1}$ . At typical biomasses before the first manipulation (edible algae  $\approx 1$  mg/L, zooplankton  $\approx 0.2$  mg/m<sup>2</sup>), net weekly growth is  $\approx 1.1$  mg/L. Weekly grazing loss is  $\approx 0.038$  mg/L, or  $\approx 3\%$  of growth. During 1986–1990, biomasses in Tuesday Lake were more similar to those of Paul Lake (edible

TABLE 2. Summary of model fits for edible phytoplankton in Paul and Tuesday Lakes. Four models (with and without density dependence; prey-dependent or ratio-dependent functional response) were analyzed for each lake. Error model (if needed), residual standard deviation, correlation of predictions and observations, and prediction standard deviation (obtained by bootstrapping) are presented for each model. Number of observations was 96 for Paul Lake and 91 for Tuesday Lake.

Lake	Density dependence	Functional response	Error model	Residual $s$	Corr.	Prediction $s$
Paul	no	PD1	ARMA (1,1)	0.0955	0.567	0.185
Paul	yes	PD1	none	0.190	0.508	...
Paul	no	RD2	ARMA (1,1)	0.0980	0.542	0.572
Paul	yes	RD2	none	0.170	0.519	...
Tues.	no	PD1	AR (1)	0.392	0.769	1.83
Tues.	yes	PD1	ARMA (1,1)	0.365	0.639	4.56
Tues.	no	RD2	none	0.257	0.784	0.810
Tues.	yes	RD2	AR (1)	0.237	0.760	3.67



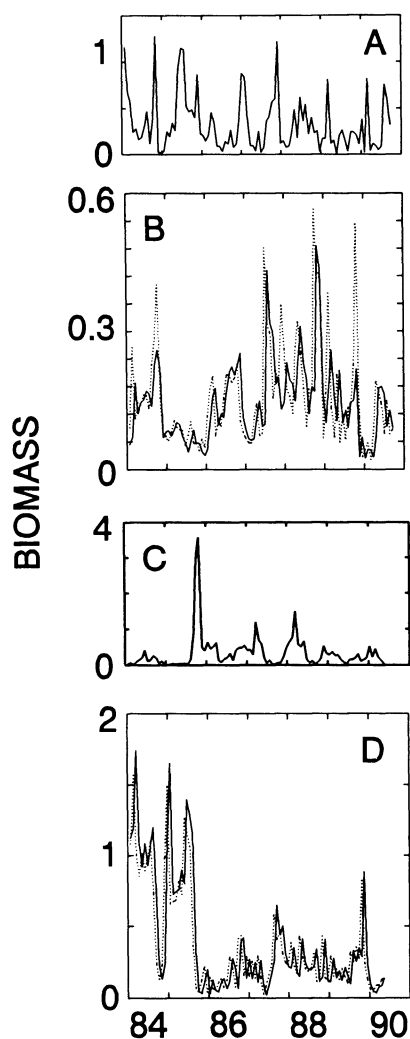


FIG. 4. Observed series and model predictions for the experimental lakes. (A) Biomass of cladoceran herbivores ( $\text{mg}/\text{m}^2$ ) in Paul Lake. (B) Biomass of edible algae ( $\text{mg}/\text{L}$ ) observed (.....) and predicted one step ahead (—) in Paul Lake. Predictions were calculated from the model with no density dependence, PD1 functional response, and ARMA(1, 1) errors. (C) Biomass of cladoceran herbivores ( $\text{mg}/\text{m}^2$ ) in Tuesday Lake. (D) Biomass of edible algae ( $\text{mg}/\text{L}$ ) observed (.....) and predicted one step ahead (—) in Tuesday Lake. Predictions were calculated from the model with no density dependence and RD2 functional response.

algae  $\approx 0.2 \text{ mg}/\text{L}$ , zooplankton  $\approx 0.5 \text{ mg}/\text{m}^2$ ). During these years, growth and loss rates in Tuesday Lake were similar to those calculated for Paul Lake: net weekly growth is  $\approx 0.21 \text{ mg}/\text{L}$ , while weekly grazing loss is  $\approx 0.044 \text{ mg}/\text{L}$ , or  $\approx 20\%$  of growth.

#### DISCUSSION

Error-in-variables problems in modeling ecological time series have been studied most thoroughly for fisheries data (Ludwig and Walters 1989, Hilborn and Walters 1992). In fisheries applications, the observation error is often poorly known. In other ecological appli-

cations, observation errors can be estimated by replicate samples. In such cases, the methods presented here are applicable.

For these models and data sets, the three loss functions we examined (Expressions 9, 11, and 12) led to identical and apparently accurate estimates of the model parameters. However, there is no guarantee that they yield reliable estimates in all situations. Bias will depend on the particular models being considered, the error distribution, and the magnitude of the errors. For applications different from those studied here, we recommend that the estimation procedures be tested by stochastic simulation.

With realistic observation errors, predator-prey models are very difficult to identify correctly. In many cases, intrasystem spatial variability will be a significant and unavoidable component of observation error. Modest increases in length of the time series may not solve this problem. For example, we repeated the analyses of Fig. 3E using time series of 100 observations instead of 50, with similar results. The underlying RD2 model was correctly identified only 32% of the time against the PD1 alternative and only 18% of the time against the PD2 alternatives.

Manipulations of the state variables substantially improve our ability to identify the underlying model. Our conclusion parallels the findings of Walters (1986) for dynamic models of fisheries.

Comparing alternative manipulations to see which ones best discriminate the models is an important question that is not addressed by this study. For example, if  $V$  and  $P$  are both perturbed to small values, but with a constant ratio, the ratio-dependent model would predict a faster recovery (proportional to  $V$ ) while the prey-dependent model would predict a slower recovery (proportional to  $VP$ ). Alternatively, if perturbation makes the ratio  $V/P$  very small while  $V$  and  $P$  remain relatively large, the prey-dependent model will predict a high predation rate while the ratio-dependent model will predict a low predation rate. The feasibility of these sorts of perturbations will depend on the ecosystem being studied. For lakes, relatively simple press perturbations of the type we studied have been performed successfully. In specific applications, simulations could be performed to determine manipulations that are most likely to be informative.

Correction for observation error affects both model identification and parameter estimates. An earlier analysis of time series from Paul Lake used conventional nonlinear maximum likelihood fits that did not correct for observation error (Carpenter et al. 1994). In that study, effects of observation error on model identification were assessed by stochastic simulation, as suggested by Hilborn and Walters (1992). Like the analysis reported here, the conventional analysis indicated that the linear functional response was the best grazing model. The conventional estimate of the grazing coefficient, per week per milligram of zooplankton per

square metre, was  $0.315 \text{ wk}^{-1} \cdot (\text{mg}/\text{m}^2)^{-1}$ , within 20% of the value of  $0.377 \text{ wk}^{-1} \cdot (\text{mg}/\text{m}^2)^{-1}$  estimated by our error-in-variables method. However, the conventional method indicated that the density-dependent equation was the best model for algal growth, while the method employed in this paper indicated that a density-independent model was more appropriate. Our simulation studies show that both model identification and parameter estimates are sensitive to observation error. We have greater confidence in the model fits reported here, which account for observation error.

The difference in the best fitting models between Paul and Tuesday Lakes is explainable by the manipulation of Tuesday Lake and may not indicate a fundamental difference between the lakes in phytoplankton-zooplankton interactions. Time series from unmanipulated Paul Lake were fit by a relatively simple linear model, while those from manipulated Tuesday Lake were fit by a more complex nonlinear model. The greater contrast in algal and zooplankton biomasses observed in Tuesday Lake supported a more complex model and allowed satisfactory estimation of more parameters.

Nonlinear functional responses may not be evident in ecological time series unless the state variables can be experimentally manipulated or otherwise perturbed. Given the modest fluctuations of unperturbed systems, prediction errors will usually be smaller for linear models, which fit about as well as nonlinear models but have fewer parameters and smaller parameter covariance matrices (Carpenter et al. 1994). This tendency to favor simpler models can be reduced by experimental manipulations of ecosystems. Even when informative manipulation is possible, some important differences among models may be indistinguishable. Therefore, some ecological mechanisms may be very difficult to distinguish at the ecosystem scale (Tilman 1989). In management applications, failure to recognize this uncertainty could cause costly errors (Walters 1986, Hilborn and Walters 1992).

For example, ratio-dependent and prey-dependent grazing models have different implications for management of lakes eutrophied by nutrient enrichment. Ratio dependence implies rapid recovery from perturbation and effective transfer of nutrients to higher trophic levels (Arditi and Ginzburg 1989). Under prey dependence, algae are suppressed or develop nuisance blooms, depending on whether large herbivorous zooplankton are present (Carpenter 1992). Managers following a ratio-dependent model would expect pelagic systems to be relatively resilient and respond smoothly to changes in nutrient loading or grazing. In contrast, managers following a prey-dependent model would expect sudden transitions in water quality associated with threshold levels of nutrient loading or grazing. These different expectations could lead to very different management strategies (Benndorf 1989). Our analyses indicate that it may be very difficult to determine which model is appropriate and that manipulations of nutri-

ent load or grazers may help to discriminate the alternatives.

In spite of these cautions, fitting models to ecological time series has significant advantages for parameter estimation and forecasting. Data can be collected at appropriate scales for the system of interest, thus bypassing the assumptions and uncertainties that arise when parameters are imported from other systems or small-scale experiments. Parameter estimates are improved by accounting for observation error. However, even modest observation errors cause substantial difficulties in model identification. Ecosystem manipulation substantially increases the probability of detecting the underlying model, reduces uncertainty in parameter estimates, and yields models capable of predicting a wider range of ecosystem states.

#### ACKNOWLEDGMENTS

We thank A. R. Ives, A. Muñoz del Río, and D. E. Schindler for helpful comments on the manuscript. Our research on modeling ecological time series is supported by the Andrew W. Mellon Foundation. The ecosystem experiments were supported by the National Science Foundation. K. L. Cottingham was supported by a NSF Predoctoral Fellowship.

#### LITERATURE CITED

- Arditi, R., and L. R. Ginzburg. 1989. Coupling in predator-prey dynamics: ratio dependence. *Journal of Theoretical Biology* **139**:311–326.
- Arditi, R., L. R. Ginzburg, and H. R. Akçakaya. 1991. Variation in plankton densities among lakes: a case for ratio-dependent predation models. *American Naturalist* **138**:1287–1296.
- Arditi, R., and H. Saiah. 1992. Empirical evidence of the role of heterogeneity in ratio-dependent consumption. *Ecology* **73**:1544–1551.
- Bates, D. M., and D. G. Watts. 1988. *Nonlinear regression analysis and its applications*. Wiley, New York, New York, USA.
- Benndorf, J. 1989. Conditions for effective biomanipulation: conclusions derived from whole-lake experiments in Europe. *Hydrobiologia* **200/201**:187–204.
- Bergquist, A. M., and S. R. Carpenter. 1986. Limnetic herbivory: effects on phytoplankton populations and primary production. *Ecology* **67**:1351–1360.
- Berryman, A. A. 1992. The origins and evolution of predator-prey theory. *Ecology* **73**:1530–1535.
- Blau, G. E., and W. B. Neely. 1975. Mathematical model building with an application to determine the distribution of Dursban insecticide added to a simulated ecosystem. *Advances in Ecological Research* **9**:133–163.
- Bohachevsky, I. O., M. E. Johnson, and M. L. Stein. 1986. Generalized simulated annealing for function optimization. *Technometrics* **28**:209–217.
- Carpenter, S. R. 1992. Destabilization of planktonic ecosystems and blooms of blue-green algae. Pages 461–482 in J. F. Kitchell, editor. *Food web management*. Springer-Verlag, New York, New York, USA.
- , editor. 1988. *Complex interactions in lake communities*. Springer-Verlag, New York, New York, USA.
- Carpenter, S. R., T. M. Frost, J. F. Kitchell, T. K. Kratz, D. W. Schindler, J. Shearer, W. G. Sprules, M. J. Vanni, and A. P. Zimmerman. 1991. Patterns of primary production and herbivory in 25 North American lake ecosystems. Pages 67–96 in J. Cole, S. Findlay, and G. Lovett, editors. *Com-*

- parative analyses of ecosystems: patterns, mechanisms, and theories. Springer-Verlag, New York, New York, USA.
- Carpenter, S. R., and J. F. Kitchell, editors. 1993. The trophic cascade in lakes. Cambridge University Press, London, England.
- Carpenter, S. R., R. C. Lathrop, and A. Muñoz-del-Rio. 1994. Comparison of dynamic models for edible phytoplankton. *Canadian Journal of Fisheries and Aquatic Sciences*, in press.
- Chandler, J. P. 1972. On an iterative procedure for estimating functions when both variables are subject to error. *Technometrics* 14:71-76.
- Clutton-Brock, M. 1967. Likelihood distributions for estimating functions when both variables are subject to error. *Technometrics* 9:261-269.
- Edwards, A. W. F. 1992. Likelihood. Expanded edition. Johns Hopkins University Press, Baltimore, Maryland, USA.
- Frost, T. M., D. L. DeAngelis, S. M. Bartell, D. J. Hall, and S. H. 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.
- Fuller, W. A. 1987. Measurement error models. John Wiley & Sons, New York, New York, USA.
- Gatto, M. 1991. Some remarks on models of plankton densities in lakes. *American Naturalist* 137:264-267.
- Ginzburg, L. R., and H. R. Akçakaya. 1992. Consequences of ratio dependent predation for steady-state properties of ecosystems. *Ecology* 73:1536-1543.
- Hanski, I. 1991. The functional response of predators: worries about scale. *Trends in Ecology and Evolution* 6:141-142.
- Hilborn, R., and C. J. Walters. 1992. Quantitative fisheries stock assessment. Chapman and Hall, New York, New York, USA.
- Holling, C. S. 1959. The components of predation as revealed by a study of small mammal predation of the European sawfly. *Canadian Entomologist* 91:293-320.
- Jassby, A. D., and T. M. Powell. 1990. Detecting changes in ecological time series. *Ecology* 71:2044-2052.
- Jassby, A. D., T. M. Powell, and C. R. Goldman. 1990. Interannual fluctuations in primary production: direct physical effects and the trophic cascade at Castle Lake, California. *Limnology and Oceanography* 35:1021-1038.
- Judge, G. G., W. E. Griffiths, R. C. Hill, H. Lutkepohl and T. C. Lee. 1985. The theory and practice of econometrics. Second edition. John Wiley & Sons, New York, New York, USA.
- Leger, C., D. N. Politis, and J. P. Romano. 1992. Bootstrap technology and applications. *Technometrics* 34:378-398.
- Leibold, M. 1990. Resource edibility and the effects of predators and productivity on the outcome of trophic interactions. *American Naturalist* 134:922-949.
- Likens, G. E., editor. 1989. Long-term studies in ecology. Springer-Verlag, New York, New York, USA.
- Ludwig, D., and C. J. Walters. 1989. A robust method for parameter estimation from catch and effort data. *Canadian Journal of Fisheries and Aquatic Sciences* 46:137-144.
- Magnuson, J. J. 1990. Long-term ecological research and the invisible present. *BioScience* 40:495-501.
- May, R. M. 1981. Models for single populations. Pages 5-29 in R. M. May, editor. *Theoretical ecology*. Second edition. Sinauer, Sunderland, Massachusetts, USA.
- McCauley, E., W. W. Murdoch, and S. Watson. 1988. Simple models and variations in plankton densities among lakes. *American Naturalist* 132:383-403.
- Morris, W. F. 1990. Problems in detecting chaotic behavior in natural populations by fitting simple discrete models. *Ecology* 71:1849-1862.
- Nelder, J. A., and R. Mead. 1964. A simplex method for function minimization. *Computer Journal* 7:308-313.
- Oksanen, L., J. Moen, and P. A. Lundberg. 1992. The time-scale problem in exploiter-victim models: does the solution lie in ratio-dependent exploitation? *American Naturalist* 140:938-960.
- Pace, M. L. 1984. Zooplankton community structure, but not biomass, influences the phosphorus-chlorophyll *a* relationship. *Canadian Journal of Fisheries and Aquatic Sciences* 41:1089-1096.
- Porter, K. G., J. Gerritsen, and J. D. Orcutt, Jr. 1982. The effect of food concentration on swimming patterns, feeding behavior, ingestion, assimilation and respiration by *Daphnia*. *Limnology and Oceanography* 27:935-949.
- Reilly, P. M., and H. Patino-Leal. 1981. A Bayesian study of the error-in-variables model. *Technometrics* 23:221-231.
- Risser, P. G., editor. 1991. Long-term ecological research: an international perspective. Wiley, New York, New York, USA.
- Rosenzweig, M. L. 1971. Paradox of enrichment: destabilization of exploitation ecosystems in ecological time. *Science* 171:385-387.
- Rudstam, L. G., R. C. Lathrop, and S. R. Carpenter. 1993. The rise and fall of a dominant planktivore: direct and indirect effects on zooplankton. *Ecology* 74:303-319.
- Sarnelle, O. 1992. Nutrient enrichment and grazer effects on phytoplankton in lakes. *Ecology* 73:551-560.
- Seber, G. A. F., and C. J. Wild. 1989. Nonlinear regression. John Wiley & Sons, New York, New York, USA.
- Slobodkin, L. B. 1992. A summary of the Special Feature and comments on its theoretical context and importance. *Ecology* 73:1564-1566.
- Sprent, P. 1990. Some history of functional and structural relationships. *Contemporary Mathematics* 112:3-15.
- Tilman, D. 1989. Ecological experimentation: strengths and conceptual problems. Pages 136-157 in G. E. Likens, editor. *Long-term studies in ecology*. Springer-Verlag, New York, New York, USA.
- Walters, C. J. 1986. Adaptive management of renewable resources. MacMillan, New York, New York, USA.
- Watson, S., and E. McCauley. 1988. Contrasting patterns of net- and nanoplankton production and biomass among lakes. *Canadian Journal of Fisheries and Aquatic Sciences* 45:915-920.
- Wei, W. W. S. 1990. Time series analysis. Addison-Wesley, New York, New York, USA.
- Whittemore, A. S. 1989. Errors-in-variables regression using Stein estimates. *American Statistician* 43:226-228.

## APPENDIX

## TIME SERIES ERROR TERMS

Terms for serially correlated errors were added to the models when residuals had significant autocorrelations or partial autocorrelations (Bates and Watts 1988, Wei 1990). In our data, the only significant autocorrelations or partial autocorrelations occurred at lag 1. Therefore, we considered autoregressive, moving average, and autoregressive moving average terms all at lag 1, referred to as AR(1), MA(1), and ARMA(1, 1) terms, respectively (Wei 1990).

Eq. 5 can be rewritten as

$$V_{t+1} = F_t + Z_t, \quad (\text{A.1})$$

where  $F_t$  contains the previous value of the response series  $V$ , the input variable  $P$  and the parameters to be estimated. Thus

$$F_t = rV_t - sV_t^2 - f[V_t, P_t] \quad (\text{A.2})$$

for the density-dependent case. For the density-independent case,  $s$  in Eq. A.2 is set to zero.  $Z_t$  is the residual series, which is serially correlated.

In the AR(1) case,

$$Z_t = \phi Z_{t-1} + \epsilon_t, \quad (\text{A.3})$$

where  $\phi$  is a parameter that accounts for the autocorrelation and  $\epsilon_t$  is the series of filtered residuals (Wei 1990). If the AR(1) correction has succeeded in removing serial correlation from the residuals, then the autocorrelation and partial autocorrelation functions of  $\epsilon_t$  will be nonsignificant. The full model is written by combining Eqs. A.1 and A.3:

$$V_{t+1} = F_t + \phi Z_{t-1} + \epsilon_t, \quad (\text{A.4})$$

where  $Z_{t-1} = V_t - F_{t-1}$ . The parameters to be estimated now include  $\phi$  and the parameters of Eq. A.2. These were estimated by minimizing Expression 11 with respect to the parameters.

For this application of Eq. 11, the residuals first defined in Eq. 10 are replaced by

$$g_t = V_{t+1} - (F_t + \phi Z_{t-1}) \quad (\text{A.5})$$

and the vector  $B_t$  is the partial derivative of  $g_t$  with respect to the observed series  $V_{t+1}$ ,  $V_t$ , and  $P_t$ .

The MA(1) and ARMA(1, 1) cases are handled in a similar way. An equation for the serially correlated errors (analogous to Eq. A.3) is combined with Eq. A.1 to yield a modified model in which errors are corrected for serial correlation (analogous to Eq. A.4). The residuals of the modified model are then used in Expression 11, which is minimized with respect to the parameters.

In the MA(1) case,

$$Z_t = \epsilon_t - \theta\epsilon_{t-1} \quad (\text{A.6})$$

(Wei 1990). The modified model is

$$V_{t+1} = F_t + \epsilon_t - \theta\epsilon_{t-1} \quad (\text{A.7})$$

where  $\epsilon_{t-1} = V_t - F_{t-1} + \theta\epsilon_{t-2}$ . The residuals used in Expression 11 are

$$g_t = V_{t+1} - (F_t - \theta\epsilon_{t-1}). \quad (\text{A.8})$$

In the ARMA(1, 1) case,

$$Z_t = \phi Z_{t-1} + \epsilon_t - \theta\epsilon_{t-1} \quad (\text{A.9})$$

(Wei 1990). The modified model is

$$V_{t+1} = F_t + \phi Z_{t-1} + \epsilon_t - \theta\epsilon_{t-1}, \quad (\text{A.10})$$

where  $Z_t = V_t - F_{t-1}$  and  $\epsilon_{t-1} = Z_{t-1} - \phi Z_{t-2} + \theta\epsilon_{t-2}$ . The residuals used in Expression 11 are

$$g_t = V_{t+1} - (F_t + \phi Z_{t-1} - \theta\epsilon_{t-1}). \quad (\text{A.11})$$