

## Comparison of First-Order Error Analysis and Monte Carlo Simulation in Time-Dependent Lake Eutrophication Models

DONALD SCAVIA

*National Oceanic and Atmospheric Administration, Great Lakes Environmental Research Laboratory  
Ann Arbor, Michigan 48104*

WILLIAM F. POWERS AND RAYMOND P. CANALE

*University of Michigan, Ann Arbor, Michigan 48109*

JENNIE L. MOODY

*National Oceanic and Atmospheric Administration, Great Lakes Environmental Research Laboratory  
Ann Arbor, Michigan 48104*

Estimates of variance for a nonlinear, seasonal food chain, nutrient cycle eutrophication model of Saginaw Bay, Lake Huron, calculated by first-order variance propagation and Monte Carlo analyses, do not always agree. A comparison of estimates of state variables indicates that Monte Carlo means are most like the measurements, whereas Monte Carlo medians are most like the deterministic model output. Best agreement between Monte Carlo and first-order estimates of both state variable values and their variances occurs when Monte Carlo output distributions are symmetric. Under these conditions, both estimates are measures of variance associated with total populations (i.e., all algae). Those distributions, however, change dramatically in time for most state variables. For asymmetric distributions, first-order variance estimates measure variability about the typical component of the total population (i.e., the typical algal species) and Monte Carlo variance estimates measure variability of the mean component (which is more reflective of the total). One must be cognizant of these differences when estimating variance associated with model projections.

### INTRODUCTION

Lake eutrophication models have become relatively common tools of resource managers. The continuum of model types range from simple nomographs [e.g., *Vollenweider*, 1969, 1975] to more complicated, time variable, phytoplankton-based food chain models [e.g., *Thomann et al.*, 1977]. Projections made with these models affect management decisions which often influence many thousands of people, economically and socially. It is therefore important to couch these projections in terms that indicate the faith one should have in them. Most models used in lake management analysis have been deterministic and thus their stochastic nature (or uncertainty) has been ignored until recently [*Reckhow*, 1979; *Lettenmaier and Richey*, 1979; *Chapra and Reckhow*, 1979]. Stochastic properties of stream quality models [mainly dissolved oxygen (DO), biological oxygen demand (BOD) models] have received more attention [e.g., *Burges and Lettenmaier*, 1975; *Moore et al.*, 1976; *Chiu*, 1978].

Models also serve a purpose other than providing projections. They can be useful tools for exploring relationships and controls within the lake [e.g., *Lehman et al.*, 1975; *Scavia*, 1979; *Scavia and Bennett*, 1980]. For example, models of lake ecosystems can be useful for comparing phosphorus-loading rates and internal cycling rates to determine the relative short-term and long-term impacts of loading alterations. Obviously, estimates of errors associated with the loading and cycling rates will provide a firmer base for their comparison.

In a subsequent paper [*Scavia et al.*, 1981], we use first-order

variance propagation to estimate overall model variance originating from uncertain initial conditions, parameters, and selected external driving forces. In that paper we compare model variance and data variance, identify the major sources of model variance, and explore model sensitivity using a correlation matrix generated from variance propagation.

The purpose of the work presented herein is to examine two methods that have been used for estimating variance associated with eutrophication model output: Monte Carlo simulation and first-order variance propagation. Herein error or variance estimates are considered to be those propagated via uncertain model initial conditions and parameters. They do not represent variance between model output and measurements.

Both methods have been used for estimating errors in water quality models. For example, *Reckhow* [1979] and *Lettenmaier and Richey* [1979] used first-order error analyses for total phosphorus, mass balance lake models; *Tiwari and Hobbie* [1976], *O'Neill and Gardner* [1979], *O'Neill et al.* [1980], and *Gardner et al.* [1980] used Monte Carlo analyses for estimating errors in more complicated food web models. Occasionally both methods have been used in the same study [e.g., *Burges and Lettenmaier*, 1975; *Montgomery et al.*, 1980]. Assumed implicitly in these latter studies, and suggested explicitly elsewhere [e.g., *Gelb*, 1974], is that Monte Carlo represents 'truth' and that it can be used to check the accuracy of approximations in the first-order variance propagation analysis. The work presented herein suggests that Monte Carlo and first-order variance propagation do not necessarily quantify the same type of variability. Suggestions based on these results indicate the conditions for which each method should be used and when they should be expected to agree.

This paper is not subject to U.S. copyright. Published in 1981 by the American Geophysical Union.

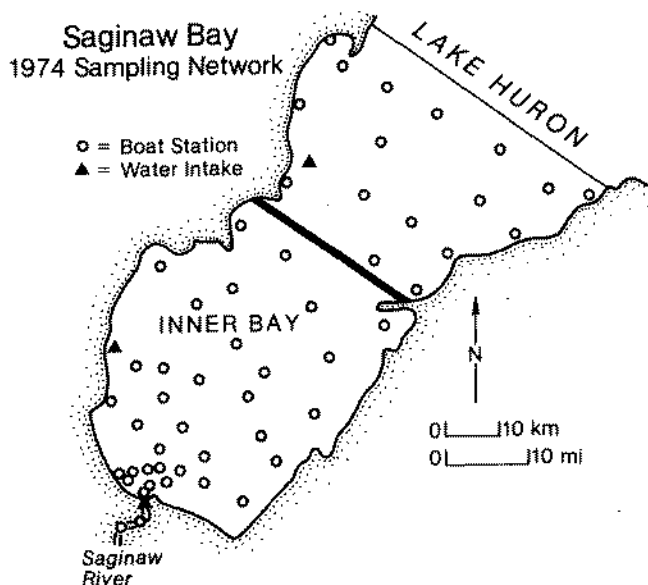


Fig. 1. Map of Saginaw Bay showing inner bay segment and sampling locations. (Redrawn from Richardson and Bierman [1976].)

#### METHODS

**The model.** The system model was developed for predicting trends of eutrophication in large lakes. Developed originally for Lake Ontario [Thomann *et al.*, 1975, 1976], it has been modified and applied in several other contexts [Canale *et al.*, 1974; DiToro and Matystik, 1976, 1980; Richardson and Bierman, 1976; Thomann *et al.*, 1977]. The original conceptualization is an eight-state-variable, nonlinear, time-dependent food chain model used herein to simulate seasonal dynamics in a vertically averaged inner portion of Saginaw Bay, Lake Huron (Figure 1). The model simulates dynamics of phytoplankton, herbivorous and carnivorous zooplankton, three nitrogen forms, and two phosphorus forms. The general interactions among state variables and examples of nonlinear formulations are illustrated conceptually in Figure 2. Detailed model equations, parameter values, and references to boundary conditions, forcing functions, and driving data can be found in Thomann *et al.* [1975], Bierman *et al.* [1980], and Scavia [1980].

The model was calibrated by manipulation of parameter values to a 1974 data set. Although some attempt was made to obtain general agreement between simulation and observation, close agreement was not obtained in all cases. Precise calibration is not necessary for the comparisons to be made below.

**Monte Carlo simulations.** For Monte Carlo simulations the following procedure was used. The model equations were solved repeatedly. Each model execution was performed with initial conditions and parameter values selected randomly from their individual distributions. From these repeated model executions, state variable means, variances, and other statistics were calculated at 4-week intervals throughout the period of simulation. The analysis was terminated after 1000 simulations, at which time state variable means and variances were converging. Further details on Monte Carlo simulations can be found in McGrath and Irving [1973].

The input sets of initial conditions and parameter values were generated from GGTRA, a subroutine [International

Mathematical and Statistical Libraries, Inc., 1977] that can produce the desired number of pseudo-random deviates from triangular distributions with specified minimum, maximum, and mode. Values of the minimum and maximum for each parameter were determined from the literature or by assumption and for each initial condition by examination of Saginaw Bay measurements.

Triangular distributions were used because available information on parameter distributions allowed no better assumption. Assuming other distributions implies further information is available. Modes were selected to correspond to constant values used in the deterministic model because they represent parameters of the typical system components, rather than the average component. This is consistent with the deterministic application. In two similar comparisons between Monte Carlo simulation and first-order variance propagation where normal distributions and means corresponding to the deterministic constants were used [Scavia, 1980], results similar to those reported below were obtained.

All initial conditions and parameters were assumed to be independent. These values and variance estimates from the 1000 generated deviates are given in Table 1. Although these quantities are surely not independent, their covariances are difficult to estimate with confidence. If they were estimated, covariances could be included in future analyses in a straightforward fashion computationally.

**First-order variance propagation.** The familiar [Cornell, 1972] first-order variance propagation formula for a scalar, time-invariant model is

$$\sigma_y^2 = \left[ \frac{\delta g(z)}{\delta z} \right]^2 \sigma_z^2 \quad (1)$$

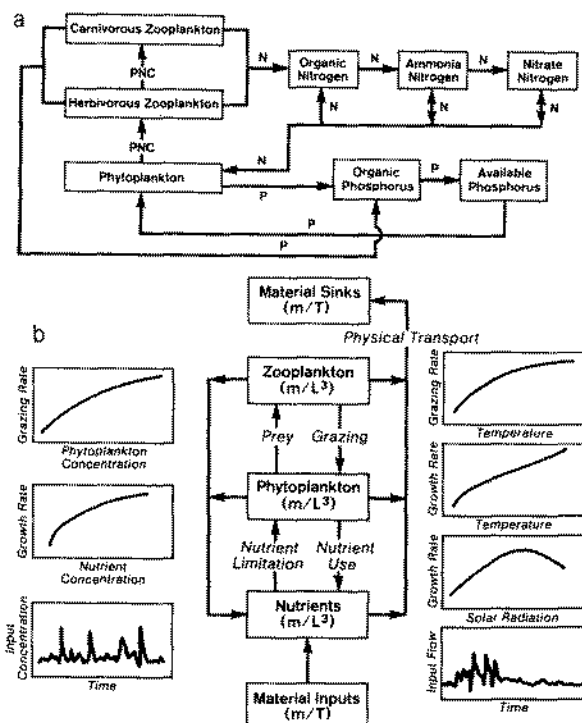


Fig. 2. (a) Conceptual model of Saginaw Bay eutrophication dynamics. P, N, and C show flow of phosphorus, nitrogen, and carbon, respectively. (Redrawn from Richardson and Bierman [1976].) (b) Conceptual illustration of inputs and nonlinear relationships in Saginaw Bay model. (Redrawn from Richardson and Bierman [1976].)

where  $\sigma_y^2$  and  $\sigma_z^2$  are the variances of  $y$ , the dependent variable, and  $z$ , the independent variable of the nonlinear algebraic model,  $y = g(z)$ . The partial derivative is usually evaluated at the mean value of  $z$ .

The more general case includes correlated independent variables ( $z_i$ ) and a system of nonlinear equations

$$y_1 = g_1(z_1, \dots, z_n, \dots, z_m)$$

$$y_i = g_i(z_1, \dots, z_n, \dots, z_m)$$

$$y_n = g_n(z_1, \dots, z_n, \dots, z_m)$$

represented by the first-order Taylor series approximation

$$y = \mathbf{A} \mathbf{Z} \quad (2)$$

where  $A_{ij} = [\delta g_i / \delta z_j]$ , is evaluated at the mean value of  $\mathbf{Z}$ .

The covariance matrix of the dependent variables ( $P_y$ ) can be estimated as a function of the independent variable covariance matrix ( $P_z$ ) from (in matrix notation)

$$P_y = \mathbf{A} P_z \mathbf{A}^T \quad (3)$$

The time propagation of covariance, from both initial state variable covariance and independent variable (or parameter) covariance, for a nonlinear time-dependent model [ $Y_{k+1} = g(Y_k)$ ] can be accomplished in a similar fashion. Given a linearized system of equations describing the transition of the dependent variables from time  $k$  to  $k+1$

$$Y_{k+1} = \Theta_k Y_k \quad (4)$$

where  $\Theta_k$  represents a first-order approximation of the dependent variables' transition matrix for a nonlinear system [ $\Theta_{ij} = \delta g_i / \delta y_j$ ], one may augment those equations with a set of equations describing the deterministic portion of the uncertain parameters ( $\beta$ )

$$\beta_{k+1} = \Delta_k \beta_k \quad (5)$$

where  $\Delta_k$ , in this case for constant parameters, is equal to the identity matrix  $\mathbf{I}$ . Combining (4) and (5) and writing them in matrix notation yields

$$X_{k+1} = \phi_k X_k$$

where

$$\mathbf{X} = \begin{bmatrix} \mathbf{Y} \\ \beta \end{bmatrix} \quad \phi = \begin{bmatrix} \delta g / \delta y & \delta g / \delta \beta \\ 0 & \mathbf{I} \end{bmatrix} \quad (6)$$

The upper left quadrant of  $\phi$  represents a linear approximation of nonlinearities among dependent state variables; the upper right quadrant represents a linear approximation of nonlinearities between the state variables and parameters. By analogy to (2) and (3) it can be seen that the covariance matrix of  $\mathbf{X}$  ( $\mathbf{P}$ ) will propagate discretely in time from  $k$  to  $k+1$  as

$$P_{k+1} = \phi_k P_k \phi_k^T$$

where  $P_k$  includes covariances of both dependent variables and uncertain parameters at time  $k$ .

Further, this discrete-time variance propagation equation can be rewritten in continuous matrix form as

$$\dot{P} = \mathbf{A} P + P \mathbf{A}^T \quad (7)$$

by employing the definition of the derivative and performing the proper limiting operations [Gelb, 1974]. The initial conditions for this matrix differential equation are the elements of  $\mathbf{P}$  at time zero, which are the variances and covariances of the vector  $\mathbf{X}$  (see (6)) at time zero. In this particular application these represent the variances and covariances of both the model initial conditions ( $\mathbf{Y}$ ) and uncertain parameters ( $\beta$ ). A more complete discussion of this approach can be found in Scavia [1980] and formal derivations can be found in Gelb [1974] and Jazwinski [1970]. In the present application, dependent variables  $\mathbf{Y}$  are those illustrated in Figure 2, and the augmented uncertain parameters  $\beta$  are those listed in Table 1. The functions  $g_i(z_1, \dots, z_n, \dots, z_m)$  are nonlinear, simultaneous, ordinary differential equations described in Thomann et al. [1975] and Scavia [1980]. Partial derivatives of those functions with respect to state variables and uncertain parameters were derived analytically. Variance estimates from this first-order, continuous, variance propagation equation, (7), with initial variances the same as those generated for the Monte Carlo simulations (Table 1), are compared to Monte Carlo estimates herein.

In a similar study, Gardner et al. [1981] compare error analysis (Monte Carlo simulation) and first-order sensitivity analysis [Tomovic and Karplus, 1963]. They conclude that the two techniques potentially identify different major error sources because the assumptions of the sensitivity analysis are violated. The two serious violations involve ignoring higher order effects by evaluating sensitivity of only one parameter at a time and ignoring higher order effects that are manifest through large parameter errors (i.e., the first-order linearization is only strictly valid for small parameter perturbations). As described briefly below and more completely in Gelb [1974], first-order variance propagation includes not only the simultaneous effects (sensitivity) of all state variables and parameters on each state variable, but also the propagation of uncertainties in parameters and state variables. Therefore, unlike sensitivity analysis, our first-order variance propagation comparisons do not ignore higher order effects caused by having several uncertain parameters.

We are still constrained to parameter perturbations within only the linear region about the model trajectories. This region is difficult to define. Gardner et al. [1981] found the validity of the linear approximation deteriorated rapidly when the coefficient of variations for the parameter of a simple but highly nonlinear function increased above 30%. They also demonstrated for a nonlinear stream ecosystem model with input errors typically near 50% (range of coefficients of variation = 10–3000%) that the linear approximation for sensitivity analysis is also questionable. Coefficients of variation in our application (Table 1) range between 10 and 40%.

The state variable model and (7) are solved simultaneously. In all cases a paired 5th-6th Runge-Kutta, variable step size algorithm, DVERK [International Mathematical and Statistical Libraries, Inc., 1977], is used.

## RESULTS AND DISCUSSION

The first-order variance propagation equation generates a time sequence of variability about the model output. In this context, model output is referred to as the deterministic solution and can be generated independently of the propagated variance. The Monte Carlo analysis generates a time sequence of variability about the mean of 1000 model simulations.

Field measurements (bay-wide means  $\pm 1$  standard devia-

TABLE 1. Input Statistics for Monte Carlo and First-Order Error Analyses of Saginaw Bay Eutrophication Model

	Min	Max	Mode	Mean	$\sigma^2$	$100*\sigma/\bar{X}$
<i>Parameters</i>						
Phytoplankton growth rate	0.301	0.857	0.650	0.627	0.013	18
Zooplankton half saturation constant	0.004	0.039	0.020	0.021	$0.52 \times 10^{-4}$	34
Zooplankton filter rate	0.011	0.104	0.080	0.073	$0.38 \times 10^{-3}$	27
Zooplankton respiration rate	0.004	0.012	0.008	0.008	$0.32 \times 10^{-3}$	22
C: Chlorophyll ratio	25.06	49.40	30.0	32.4	28.26	16
Phytoplankton half saturation constant	0.003	0.007	0.005	0.005	$0.77 \times 10^{-6}$	18
P: Chlorophyll ratio	0.403	1.47	0.601	0.714	$0.56 \times 10^{-1}$	33
<i>Initial Conditions</i>						
Phytoplankton (mg Chl l <sup>-1</sup> )	0.004	0.019	0.009	0.010	$0.89 \times 10^{-5}$	30
Herbivores (mg C l <sup>-1</sup> )	0.003	0.012	0.008	0.008	$0.32 \times 10^{-5}$	24
Organic N (mg N l <sup>-1</sup> )	0.079	0.198	0.140	0.140	$0.59 \times 10^{-3}$	17
Ammonia (mg N l <sup>-1</sup> )	0.006	0.056	0.031	0.031	$0.10 \times 10^{-3}$	33
Nitrate (mg N l <sup>-1</sup> )	0.225	1.963	0.752	0.867	0.13	42
Organic P (mg P l <sup>-1</sup> )	0.001	0.039	0.020	0.020	$0.61 \times 10^{-4}$	39
Available P (mg P l <sup>-1</sup> )	0.002	0.020	0.008	0.009	$0.13 \times 10^{-4}$	39
Carnivores (mg C l <sup>-1</sup> )	0.015	0.049	0.032	0.032	$0.47 \times 10^{-4}$	21

tion), deterministic solutions, stochastic (Monte Carlo) means, and stochastic medians are illustrated in Figure 3 for four important state variables. The deterministic solution tracks the Monte Carlo medians more closely than the mean. This is similar to results obtained by *Burges and Lettenmaier* [1975] for a simplified Streeter-Phelps equation for DO and BOD. In their simulations the deterministic model diverged from the Monte Carlo mean and tracked the median as the stochastic model output became skewed. Output distributions in the present work were almost always skewed (see below).

The stochastic means are better estimates of both the magnitude and dynamics of the measurements than are the deterministic solutions or the medians. *Tiwari and Hobbie* [1976] and *Tiwari et al.* [1978] found that for nonlinear ecological models the range of Monte Carlo model solutions agreed well with the range of measurements and also that deterministic solutions represented only a typical measurement trajectory. In general, a deterministic model solution and medians track a typical species of phytoplankton or zooplankton and its effects on nutrients, whereas the Monte Carlo mean trajectory represents a composite average of many different species (i.e., Monte Carlo simulations use different parameter values and initial conditions, each representing theoretically different species). Typical species are not always representative of the total assemblage, especially in regard to their effects on the rest of the ecosystem. Field measurements represent a time-slice of the composite assemblage of populations and their effects. Because the mean tracks the total quantity, one should, in general, expect the Monte Carlo mean to better represent the measurements.

Variance estimates from the two approaches are qualitatively similar. In most cases they represent similar spread about the state variable estimates (Figure 4). There are, however, significant differences between the two estimates for some variables at certain times. Variance estimates from the two methods also diverged for a DO-BOD model analysis [*Burges and Lettenmaier*, 1975] as travel time downstream increased.

The differences can be attributed to four causes: (1) deterministic and stochastic mean trajectories are different, (2)

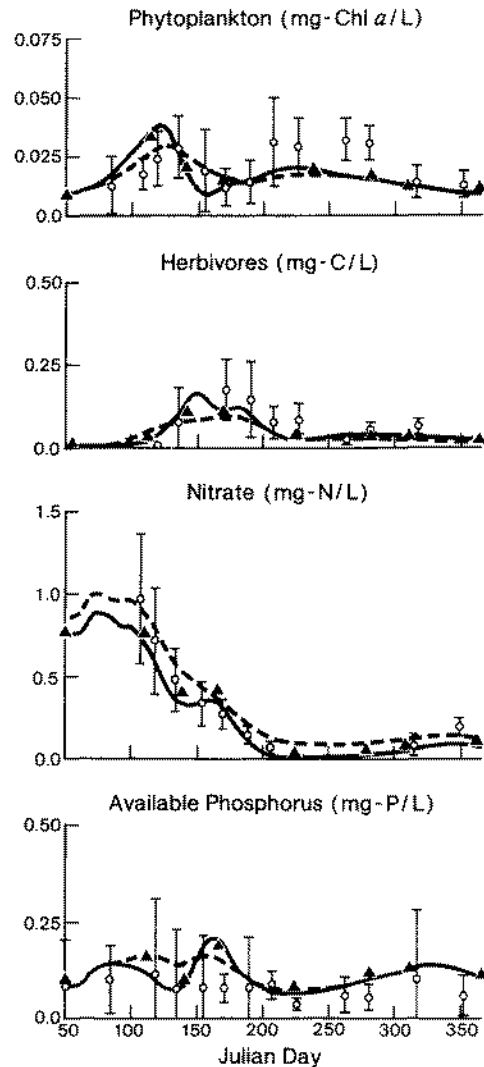


Fig. 3. Plots of deterministic model solutions (solid curve), Monte Carlo means (broken curve), medians (closed triangles), and bay-wide mean  $\pm 1$  standard deviation measurements (open circles with error bars) versus time for four model state variables.

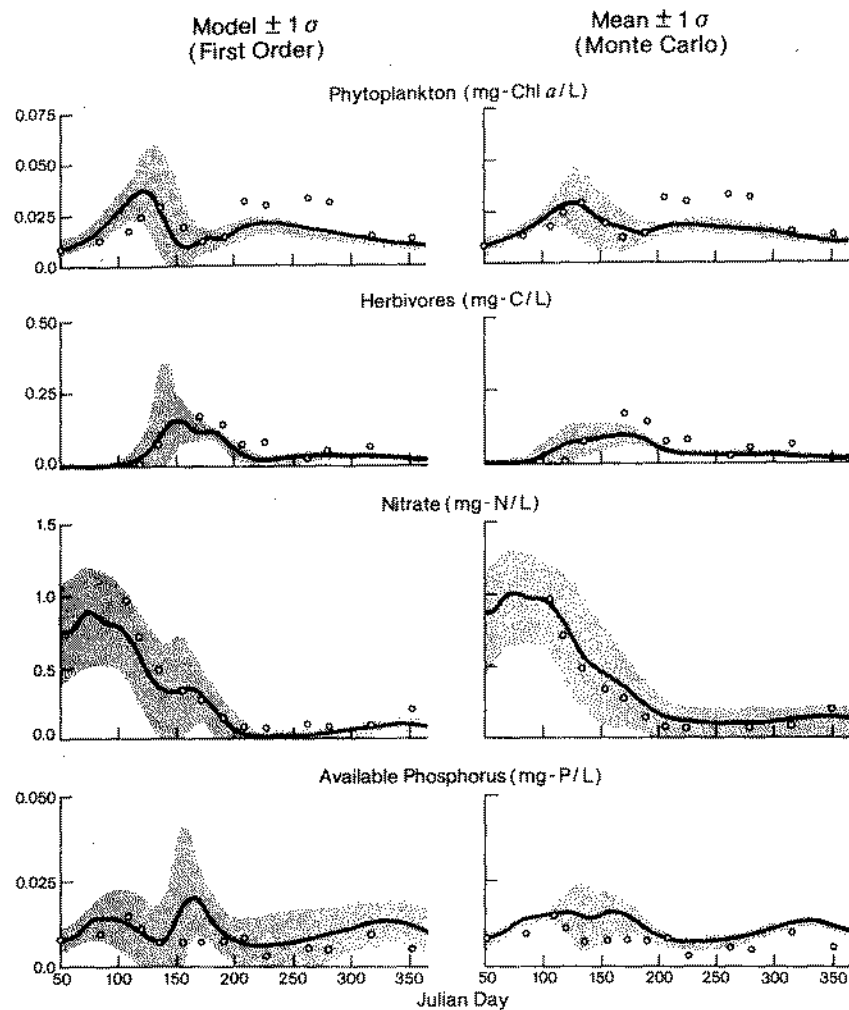


Fig. 4. Plots of deterministic model solution  $\pm 1$  standard deviation band from first-order error propagation (left panel) and Monte Carlo mean  $\pm 1$  standard deviation (right panel) versus time for four state variables. Data (open circles) are bay-wide means.

first-order error propagation employs a first-order linearization of the model, whereas Monte Carlo analysis uses the fully nonlinear model, (3) ambiguous statistics were generated from skewed Monte Carlo populations, and (4) there is a fundamental difference in the interpretation of variances from the two approaches.

Variance estimates are biased by the mean value because variance is a dimensional quantity. To reduce the effect of differing mean trajectories (Figure 3), coefficients of variation, representing percent deviation ( $100\sigma/\bar{X}$ ) from the mean, were calculated (Figure 5). Portraying variability in this way illustrates that these estimates of variance, even when scaled by the mean, can differ considerably, and thus the effect of differing mean trajectories alone is most likely not a major effect here.

Monte Carlo analysis involves repeated solution of the nonlinear model. Derivation of the first-order error propagation equations involves replacing the nonlinear model with its first-order Taylor series approximation [Gelb, 1974], and thus discrepancies between the two estimates could be caused by an inadequacy of this linear approximation. In order to examine this, comparisons were made among Monte Carlo, first-order, and second-order variance estimates. Because the second-

order propagation equation [Athans *et al.*, 1968] involves second partial derivative matrices, which are somewhat cumbersome to handle, a smaller system model that includes most of the salient nonlinear features of the eutrophication model was used for comparison. This smaller system approximates an unforced phytoplankton-zooplankton-nutrient system in a completely mixed volume with uncertain initial conditions, half saturation constant relating phytoplankton growth to nutrient concentration, and zooplankton excretion rate [Scavia, 1980]. About 5 consecutive years were simulated. While, in general, second-order propagation estimates were closer to Monte Carlo estimates than were first-order estimates, the agreement was not improved significantly. While including one higher order term in the Taylor series expansion does not guarantee a more accurate approximation, the fact that, in this case, second-order estimates were not far from first-order estimates suggests that the differences between Monte Carlo and first-order propagation variance estimates in the larger eutrophication model may not be due to any great extent to inadequacy of the first-order linearization.

The third potential cause of discrepancy between estimates of variance from the two methods is related to generation of ambiguous Monte Carlo statistics for seriously skewed popu-

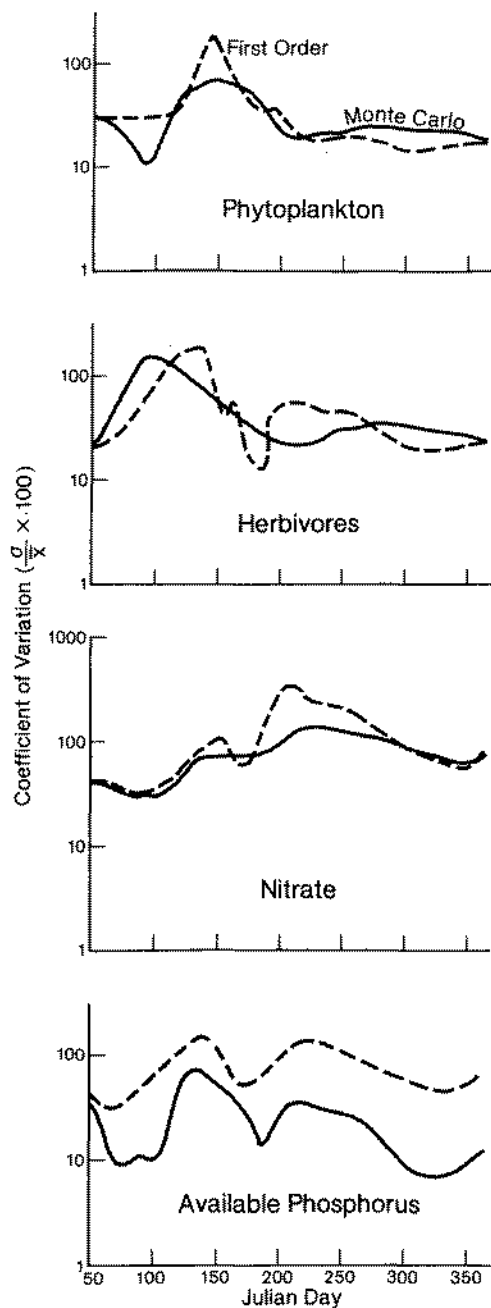


Fig. 5. Plots of coefficient of variation ( $100\sigma/\bar{X}$ ) from first-order and Monte Carlo analysis versus time for four state variables.  $\bar{X}$  is the deterministic model solution for first-order analysis, and is the mean trajectory for Monte Carlo analysis.

lations. When populations are strongly skewed, as is often the case here, variance calculated in the usual way can be an ambiguous measure of true variability. This is especially true for models that produce results indicating the possibility of bifurcation. Distribution of certain variables from Monte Carlo analysis of this model appear strongly bimodal at times (Figure 6). Under certain conditions during the Monte Carlo simulation, state variables accumulate for a given time slice at either high or low values with relatively few intermediate values. Under these conditions, variance estimates are not very useful.

Variance estimates from the two methods also diverged for

a DO-BOD model as Monte Carlo output became skewed [Burgess and Lettenmaier, 1975]. A better measure of spread in these cases is the interquartile range [Sokal and Rohlf, 1969] or, graphically, plotting the first ( $Q_1$ ), second (median), and third ( $Q_3$ ) quartiles of the population. By definition, the interval between  $Q_1$  and  $Q_3$  includes 50% of the population. This use of quartiles better represents the spread in the central portion of the skewed distributions, but it ignores the distribution tails. Plotted on Figure 7 are  $\pm 1$  standard deviation region about the deterministic solution from the first-order analysis,  $Q_1$ ,  $Q_3$ , the Monte Carlo median, and the  $\pm 1$  standard deviation region about the Monte Carlo mean. In most cases the first and third quartiles coincide generally with the two-standard deviation region about the Monte Carlo mean, illustrating that errors due to biased variance calculations for skewed population distributions are probably not serious in this case.

The fourth proposal for explaining the differences in variance estimates is related to a fundamental difference in the two approaches. In deriving the first-order error propagation equation, the mean squared deviation from the state variable true value is examined. It follows directly from use of the first-order Taylor series approximation for the nonlinear model [Gelb, 1974, section 6.1] that, in this case, 'true values' are the deterministic model solutions. Since, as discussed previously, the deterministic model attempts to track entire biological assemblages with a typical representative, variance estimates from the first-order analysis represent variability around that representative or the uncertainty associated with representing dynamics of the composite assemblage with dynamics of a typical component of the assemblage. Variance estimates from Monte Carlo analyses measure the spread of the total population about the mean trajectory. The mean trajectory estimates the dynamics of the total population; thus Monte Carlo variance estimates represent the variability or spread of individual components in the assemblage around the mean component (representing the total), rather than around the typical component, as in the first-order case. Thus variances, as calculated by the two techniques, are not necessarily the same estimators. They estimate the same property only when the deterministic solution is a good representative of the total population or, in this case, when it closely resembles the Monte Carlo mean trajectory. This is demonstrated in the latter half of the annual simulations reported herein (Figures 3 and 7) and in the shorter travel times in the work by Burgess and Lettenmaier [1975], where state variable and variance estimates agree more closely.

#### CONCLUSIONS

The above analysis reveals that variance estimates from first-order error propagation and Monte Carlo analysis are generally similar but have some significant differences (Figure 4). Our analysis suggests that these differences could not be resolved completely by (1) differences in estimates of the state variables (Figure 5), (2) errors in the first-order approximations, or (3) errors due to biased variance calculations for the skewed Monte Carlo output distributions (Figure 7). Furthermore, we suggest that the discrepancy is most likely caused by a fundamental difference in the interpretation of first-order and Monte Carlo variances.

Selection of one approach over the other depends on the type of model solution and associated variability required. If one is concerned with simulating the interactions of a heterogeneous population of organisms and their surrounding envi-

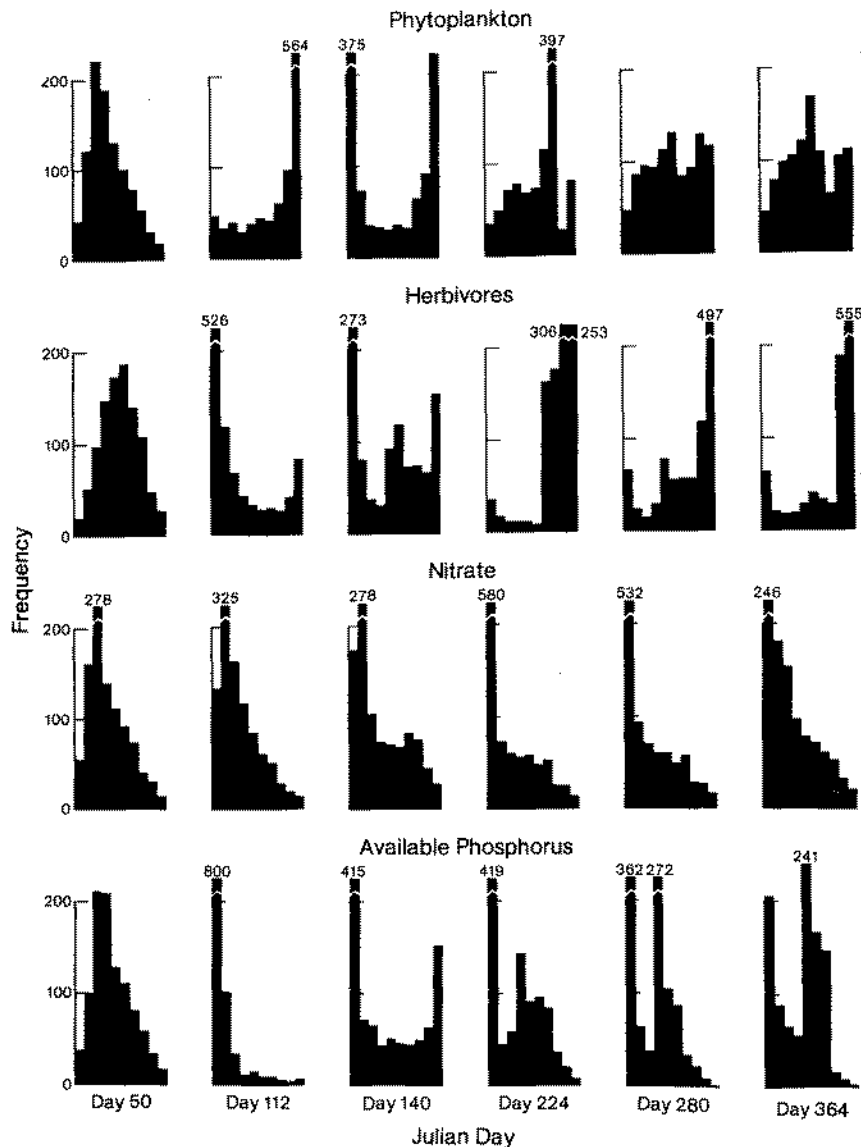


Fig. 6. Frequency distributions of Monte Carlo output for four state variables at selected time slices during annual simulation. Day 50 distributions are initial distributions. 1000 cases were used.

ronment, simulation of a single 'typical' component of that population is likely to be inadequate because the entire population is generally not well represented by that typical component alone. Simulation of typical components has, however, traditionally been the most common application. A Monte Carlo approach to the problem, using repeated simulations of the same deterministic model with different initial conditions and parameters, will provide better estimates of the total population and estimates of variance consistent with the above concern. In lieu of the full Monte Carlo approach, simultaneous simulation of several representative components of the total population would also provide a better estimate of the total, provided each component could be represented adequately by a typical representative. Under this condition, the first-order analysis could then be used for variance estimates.

Monte Carlo analysis may not be practical for long-term projections with detailed eutrophication models. However, if a deterministic model is to be used for projections, the error as-

sociated with those projections should be estimated, and the first-order error propagation procedure is an appropriate technique. Variance estimates from that procedure represent variability around the typical representative and thus can be used to assign confidence limits in the projection. Although these estimates of variance will differ from those of a more fully stochastic approach (i.e., Monte Carlo) for reasons discussed above, the differences, in this application, are not great, and the estimates provide general guides as to prediction confidence.

To explore the relative impacts of various error sources on model output variance, one often repeats the error analysis under differing input error scenarios. This procedure amplifies the required number of Monte Carlo simulations for even single-year studies. For this type of analysis the first-order propagation equation has to be solved only once for each scenario. It also provides direct estimates of the sensitivity of model output variance to particular uncertain inputs [see Scavia et al., 1981].

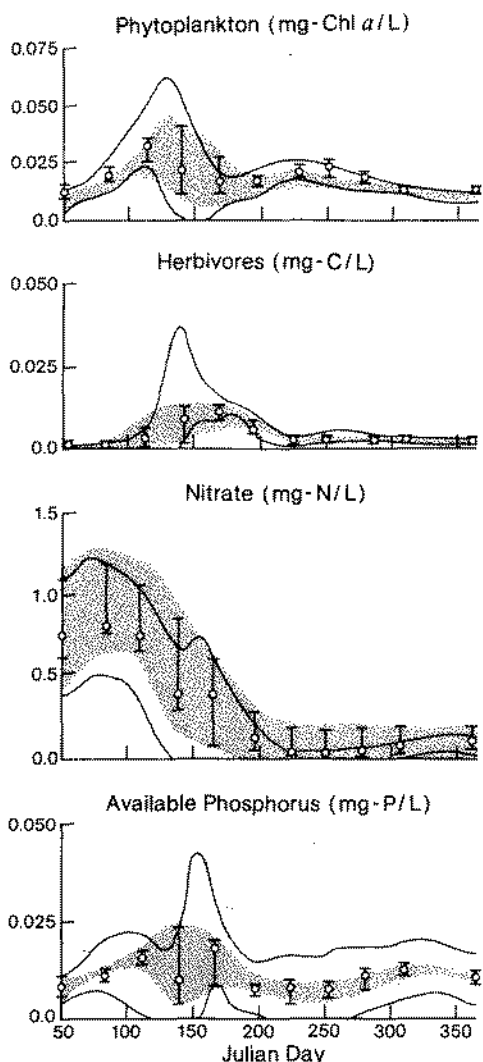


Fig. 7. Plots of deterministic model  $\pm 1$  standard deviation band from first-order analysis (open region), Monte Carlo median and first and third quartiles (open circles with error bars), and Monte Carlo mean  $\pm 1$  standard deviation band (shaded region) versus time for four state variables.

In conclusion, the Monte Carlo analysis has the advantages of estimating variability about the population mean (usually the quantity of interest) and of providing output frequency distributions. It has the disadvantage of often carrying a high computational burden for both long-term predictions and examination of relative error sources in relatively complex models. The first-order variance propagation analysis can also provide error estimates and direct estimates of model sensitivity; however, it must be recognized that those estimates reflect variability about typical components of the modeled population and that those components do not necessarily represent the dynamics of the total population or its mean.

**Acknowledgments.** We would like to acknowledge A. Robertson and D. M. DiToro for their support and helpful discussions. Comments on an earlier draft of this manuscript by R. V. O'Neill, H. A. Vanderploeg, L. R. Herche, J. M. Kelley, and two anonymous reviewers greatly improved the paper. Great Lakes Environmental Research Laboratory Contribution 234.

## REFERENCES

- Athans, M., R. P. Wishner, and A. Bertolini, Suboptimal state estimation for continuous-time nonlinear systems from discrete noisy measurements, *IEEE Trans. Autom. Control*, 504-514, Oct. 1968.
- Bierman, V. J., Jr., D. M. Dolan, E. F. Stoermer, J. E. Gannon, and V. E. Smith, The development and calibration of a multi-class, internal pool, phytoplankton model for Saginaw Bay, Lake Huron, GLERS Rep. 33, 126 pp., Great Lakes Basin Comm., Ann Arbor, Mich., 1980.
- Burges, S. J., and D. P. Lettenmaier, Probabilistic methods in stream quality management, *Water Resour. Bull.*, 11, 115-130, 1975.
- Canale, R. P., D. F. Hineman, and S. Nachippan, A biological production model for Grant Traverse Bay, *Tech. Rep. 37*, 116 pp., Univ. of Mich. Sea Grant Program, Ann Arbor, Mich., 1974.
- Chapra, S. C., and K. H. Reckhow, Expressing the phosphorus loading concept in probabilistic terms, *J. Fish. Res. Board Can.*, 36, 225-229, 1979.
- Chiu, C. L. (Ed.), *Applications of Kalman Filter to Hydrology, Hydraulics, and Water Resources*, 783 pp., Stochastic Hydraulic Program, Department of Civil Engineering, University of Pittsburgh, Pittsburgh, Pa., 1978.
- Cornell, C. A., First-order analysis of model and parameter uncertainty, paper presented at the International Symposium on Uncertainties in Hydrologic and Water Resource Systems, Univ. of Ariz., Tucson, 1972.
- DiToro, D. M., and W. Matystik, Phytoplankton biomass model of Lake Huron and Saginaw Bay, *U.S. Environ. Prot. Agency Off. Res. Dev. Rep.*, EPA 600/9-76-016, 614-618, 1976.
- DiToro, D. M., and W. Matystik, Mathematical models of water quality in large lakes, 1, Lake Huron and Saginaw Bay model development, verification, and simulations, *U.S. Environ. Prot. Agency Off. Res. Dev. Rep.*, EPA-600/3-80-056, 1980.
- Gardner, R. H., R. V. O'Neill, J. B. Mankin, and D. Kumar, Comparative error analysis of six predator-prey models, *Ecology*, 61, 323-332, 1980.
- Gardner, R. H., R. V. O'Neill, J. B. Mankin, and J. H. Carney, A comparison of sensitivity analysis and error analysis based on a stream ecosystem model, *Ecol. Modell.*, in press, 1981.
- Gelb, A. (Ed.), *Applied Optimal Estimation*, 374 pp., Massachusetts Institute of Technology Press, Cambridge, Mass., 1974.
- International Mathematical and Statistical Libraries, Inc., Subroutine DVERK, Reference Manual (Library 3), vol. 1, Houston, Tex., 1977.
- Jazwinski, A. H., *Stochastic Processes and Filtering Theory*, 376 pp., Academic, New York, 1970.
- Lehman, T. D., D. B. Botkin, and G. E. Likens, The assumptions and rationales of a computer model of phytoplankton population dynamics, *Limnol. Oceanogr.*, 20, 343-364, 1975.
- Lettenmaier, D. P., and J. E. Richey, Use of first-order analysis in estimating mass balance errors and planning sampling activities, in *Theoretical Systems Ecology: Advances and Case Studies*, edited by E. Halfon, pp. 80-106, Academic, New York, 1979.
- McGrath, E. J., and D. C. Irving, Techniques for efficient Monte Carlo simulation, Office of Naval Res. Rep. AD 762 721-723 (3 volumes), 408 pp., Natl. Tech. Inform. Serv., Springfield, Va., 1973.
- Montgomery, R. H., V. D. Lee, and K. H. Reckhow, A comparison of uncertainty analysis techniques: First-order analysis U.S. Monte Carlo simulation, paper presented at 1980 Conference on Great Lakes Research, Int. Assoc. for Great Lakes Res., Queens College, Kingston, Ont., 1980.
- Moore, S. F., G. C. Dandy, and R. J. Delucia, Describing variance with a simple water quality model and hypothetical sampling programs, *Water Resour. Res.*, 12, 795-804, 1976.
- O'Neill, R. V., and R. H. Gardner, Sources of error in ecological models, in *Methodology in Systems Modelling and Simulation*, edited by B. P. Zeigler, M. S. Elzas, G. J. Klir, and T. I. Oren, pp. 447-463, North-Holland, Amsterdam, 1979.
- O'Neill, R. V., R. H. Gardner, and J. B. Mankin, Analysis of parameter error in a nonlinear model, *Ecol. Modell.*, 8, 297-312, 1980.
- Reckhow, K. H., Empirical lake models for phosphorus development, applications, limitations, and uncertainty, in *Perspectives on Lake Ecosystem Modeling*, edited by D. Scavia and A. Robertson, pp. 183-222, Ann Arbor Science, Ann Arbor, Mich., 1979.
- Richardson, W. L., and V. J. Bierman, A mathematical model of pollutant cause and effect in Saginaw Bay, Lake Huron, U.S. Environ.



- Prot. Agency Off. Res. Dev. Rep., EPA-600/3-76-079, 138-158, 1976.
- Scavia, D., Examination of phosphorus cycling and control of phytoplankton dynamics in Lake Ontario with an ecological model, *J. Fish. Res. Board Can.*, 36, 1336-1346, 1979.
- Scavia, D., Uncertainty analysis of a lake eutrophication model, Ph.D. dissertation, 159 pp., Univ. of Mich., Ann Arbor, 1980.
- Scavia, D., and J. R. Bennett, The spring transition period in Lake Ontario—A numerical study of the causes of the large biological and chemical gradients, *Can J. Fish. Aquat. Sci.*, 37, 823-833, 1980.
- Scavia, D., R. P. Canale, W. F. Powers, and J. L. Moody, Variance estimates for a dynamic eutrophication model of Saginaw Bay, Lake Huron, *Water Resour. Res.*, this issue.
- Sokal, R. R., and F. J. Rohlf, *Biometry*, W. H. Freeman, 776 pp., San Francisco, Calif., 1969.
- Thomann, R. V., D. M. DiToro, R. P. Winfield, and D. J. O'Connor, Mathematical modeling of phytoplankton in Lake Ontario, 1, Model development and verification, *U.S. Environ. Prot. Agency Off. Res. Dev. Rep.*, EPA-660/3-75-005, 1975.
- Thomann, R. V., R. P. Winfield, D. M. DiToro, and D. J. O'Conner, Mathematical modeling of phytoplankton in Lake Ontario, 2, Simulations Using Lake I Model, *U.S. Environ. Prot. Agency Off. Res. Dev. Rep.*, EPA-600/3-76-065, 1976.
- Thomann, R. V., R. P. Winfield, and D. S. Szmuski, Estimated response of Lake Ontario phytoplankton biomass to varying nutrient levels, *J. Great Lakes Res.*, 3, 123-131, 1977.
- Tiwari, J. L., and J. E. Hobbie, Random differential equations as models of ecosystems: Monte Carlo simulation-approach, *Math. Biosci.*, 28, 25-44, 1976.
- Tiwari, J. L., J. E. Hobbie, J. P. Reed, D. W. Stanley, and M. C. Miller, Some stochastic differential equation models of an aquatic ecosystem, *Ecol. Modell.*, 4, 3-27, 1978.
- Tomovic, R., and W. J. Karplus, *Sensitivity Analysis of Dynamic Systems*, McGraw-Hill, New York, 1963.
- Vollenweider, R. A., Possibilities and limits of elementary models concerning the budget of substances in lakes (in German), *Arch. Hydrobiol.*, 66, 1-36, 1969.
- Vollenweider, R. A., Input-output models with special reference to the phosphorus loading concept in limnology, *Schweiz. Z. Hydrol.*, 37, 53-84, 1975.

(Received September 9, 1980;  
revised February 2, 1981;  
accepted February 24, 1981.)