Progress Report

on

Water Research Center Project

entitled

TEMPORAL SAMPLING STRATEGY FOR MEASURING POLLUTANT LEVELS IN WATER SYSTEMS

E-011

by

Robert J. Mulholland
School of Electrical Engineering
Oklahoma State University

July 31, 1982

PROJECT SUMMARY

This project seeks to apply the results of research sponsored by the U. S. Environmental Protection Agency, through grant CR806330-01-2, to water pollution problems of the State of Oklahoma. The purpose of the research was to develop and test an experimental design for laboratory systems which can be used as a screening tool for predicting the fate and effect of toxic materials in aquatic ecosystems. For example, research has disclosed a method for computing bounds on the time required for a pollutant to flush from a system after the source of pollution has been removed. The bounds depend on the system turnovers and the minimum level of pollutant detection using prescribed measurement techniques.

Data obtained from EPA laboratory experiments with the pesticides methoxychlor and methyl parathion applied to both flowthrough and batch aquatic microcosms have been analyzed. The time scale required to characterize the degradation of these pesticides in water is of the order of 50 days. However, it has been found that a very rapid rate process takes place between the water and surface of the sediments. An estimate for the sampling period required for the characterization of the exchange between water and sediments for methyl parathion is about 6 hours. Estimates such as these on the dynamics of a water system serve to define a window for acceptable sampling periods. The sample period window defined has been observed for both simulated and real experimental data. Furthermore, models fit to measured data have been used to verify the experimental design and sampling strategy employed. In one particular case, a model was used to discover a flow in the design of the EPA microcosm experiments. As a result, these observations have caused the re-design of experiments at the Athens (Georgia) EPA Laboratory.

I. PROJECT OBJECTIVE

The purpose of this research project was to apply systems analysis to the study of toxic materials in water systems. The investigation employed a methodology which has become quite useful for the following functions: designing laboratory experiments, processing and interpreting the results of experiments, developing model structures for systems based on experimental data, testing the predictive capabilities of evaluative models, and analyzing measurement errors.

The research will continue to refine and apply an experimental design for the analysis and evaluation of models of environmental transport, transformation, degradation, and fate of toxic materials in water systems. It is believed the methodology under study can be extended to a procedure by which models are tested against laboratory data and refined until they realistically represent perturbed natural systems. The project concentrated on applying a temporal sampling theory for measuring the movement of potentially toxic materials through models which have parameters characteristic of Oklahoma water systems. Finally, a way was sought for constructing a periodic sampling time schedule for verifying the removal or reduction in level of toxic inputs, and the subsequent prediction of the recovery time for an aquatic system.

II. RELEVANCE OF RESEARCH

More effort is needed to protect natural water resources and prevent adverse health and ecological effects from the release of potentially toxic materials, of both natural and human origin, into the environment. Toward that objective it is desirable to be able to predict the transport, transformation, degradation, fate, and effect of pollutants and other materials which impact upon water systems. After fate and effect are

determined, the regulation of the impact of toxic materials and the potentially safe levels of exposure can be established. However, the ability to enforce regulations depends critically on the accuracy with which measurements of a system can be made. One way to determine this question of accuracy is to use systems analysis and mathematical modeling to assess the inherent errors present in a measurement scheme and then compute how they affect the monitoring of pollutant levels.

The idea that systems analysis has application to water research is not new. The new concept is that systems analysis and mathematical modeling can be used to develop a quantitative theory for measuring the important parameters in water systems. The idea is to use systems analysis, much in the same way statistics is now used, to design experiments. One important question which can be answered by using this methodology is how frequently should the variables of an system be sampled in time so as to determine the dynamics of the movement of toxic materials through the system. For laboratory systems the problem of determining a proper sampling time interval has been solved. Future studies will begin the transfer of this laboratory information to field studies of natural water systems. Of particular relevance is the plan to determine the applicability of the sampling theory to models of pollutant and toxic material movement in water systems characteristic of Oklahoma. Research of this type can be viewed as a first step toward the construction of a powerful quantitative tool for the protection, conservation, and management of water resources.

III. METHODOLOGY

Prony's (1795) method is a very old technique for fitting exponential functions of time to discrete data obtained by periodically sampling a

continuous function of time. Perhaps the best description of the method is provided by Hildebrand (1956), however many other books in the area of numerical analysis tend to overlook the technique or dismiss it as being unpractical and prone to large errors. It is suspected that a lack of an adequate sampling theory may be responsible for this systematic neglect of a technique which could become a useful method for data analysis. Indeed, it can be shown that Prony's method fails to provide meaningful results when the sampling time is too long to capture the required system dynamics.

Because the solution of a compartment model is a sum of exponential time functions, Prony's method seems most appropriate for the data analysis of the measurements provided by EPA. Therefore, theoretical research was conducted in order to determine the data requirements needed for the application of Prony's method to compartment models. This research has yielded many extremely interesting results which have been described in the paper by Mulholland, Crittenden, and Weidner (1980) and in Mulholland (1981). The main result presented in this report is a sampling strategy for periodic data collection. However, of equal importance is the discussion of relationship of Prony's method to that of regression analysis for model identification. This research is a first step toward the development of an experimental design for analyzing the movement of toxic materials through laboratory ecosystems.

All microcosm experiments considered by this research project were conducted at the EPA Environmental Research Laboratory in Athens, GA. The research objectives of this project have been tested against the data obtained from these experiments. The findings of the data analysis performed is summarized in that which follows. Complete details of the EPA project are described by Mulholland (1981).

IV. RESULTS

Under US EPA cooperative research agreement No. CR-806330, data on the movement of toxic materials through controlled aquatic microcosms has been obtained. The toxicants used in these experiments are the pesticides Methoxychlor and Methyl Parathion. The data collected measure the concentrations of the pesticide and its degradation products in water as a time series of discrete values. Data sets for two different experiments exist. One concerns a flow-through system perturbed by a constant infusion of pesticide, while the other is a flask or batch microcosm into which a single dose of pesticide has been introduced. The data for both of these experiments were analyzed during the course of the research project.

Flask Study

Data obtained from the US EPA Environmental Research Laboratory in Athens, Georgia have been analyzed by using a least-squares approach based upon Prony's Method. These data relate to batch experiments performed with steady-state flask microcosms exposed to a single dose of the pesticide methyl parathion. The variable

$$x(t) = p_1 e^{\lambda_1 t} + p_2 e^{\lambda_2 t} \qquad (1)$$

represents the degradation products of the pesticide in the water compartment. The identification of the chemical rate processes taking place in the microcosm depends upon estimating the coefficients p_1 and p_2 and the exponents λ_1 and λ_2 by fitting equation (1) to the measured data.

The scientists at EPA used a standard nonlinear curve fit to obtain the following estimates using all the measured data:

$$p_1 = 0.252 \qquad \lambda_1 = -0.062$$

$$p_2 = -0.252 \qquad \lambda_2 = -0.402$$

The estimation technique used by EPA was similar to that of Muir (1980). This result should be compared with the fits to the data obtained using the least squares Prony's method for sample periods of 1, 2, 3, and 4 days which are summarized in Table 1.

As is clear from comparing the two types of fits to the data an optimum sample period of somewhere near 2.5 days should be used for future experiments. For sample periods faster or slower than this errors occur. The faster sampling times emphasize the system noise because the inherent processes are not rapidly changing, while the slower sample periods do not adequately capture the system dynamics.

Several results are evident from the analysis of the data provided by EPA. First, the design of future batch experiments involving the pesticide methyl parathion, or similar chemicals, should be carried out with a sample period of between 2 and 3 days. This sample period should be sufficient for all microcosm compartments because all rate constants of the processes involved appear in the water compartment for degradation products.

Second, in field studies involving the release of a pesticide like methyl parathion, the impact on the aquatic ecosystem can be expected to be similar to the laboratory study. In this case, a time estimate for the pesticide to flush from the system can be obtained by solving for t in (1) so that

$$x(t) = 10^{-d} \qquad (2)$$

that is the time t required for the pesticide and its degradation products to be reduced to the minimum (decimal) level of detection, 10^{-d}, where d is the precision of the measuring instruments. For example, let d = 2 then x(t) = 0.01 and solve (1) for t. This gives t = 52 days as the time required for x(t) to reach 0.01 nC/ml.

TABLE 1. LEAST-SQUARE PRONY'S METHOD

Fit to Data on Methyl Parathion

Provided by EPA

Sample Period τ	No. of Data Points	λ_1	λ_2	p_1	p_2
DAYS	M	$DAYS^{-1}$	$DAYS^{-1}$	$nCi\ ml^{-1}$	$nCi\ ml^{-1}$
1	24	0.0030	-3.7368	0.1021	-0.1029
2	12	-0.0664	-0.3935	0.2574	-0.2578
3	8	-0.0590	-0.4504	0.2390	-0.2391
4	6	-0.0440	-0.7693	0.1868	-0.1869

Third, the maximum concentration for x(t), as given by (1), is easily computed from

$$d\,x(t)/dt = 0 \tag{3}$$

which results in

$$t_m = [\ln (p_1\lambda_1/p_2\lambda_2)]/(\lambda_1 - \lambda_2) \tag{4}$$

as the time at which maximum is achieved. From (4), t_m = 5.5 days. And, the maximum concentration, computed from (1) is $x(t_m) = x(5.5) = 0.151$ nc/ml. Thus, from previous calculations, it is clear that after 50 days more than 90% of the toxic material has been degraded.

Monte Carlo simulation

In order to strengthen the results presented for the existence of an optimal sample period, a Monte Carlo experiment was performed in which exact solutions of

$$\begin{aligned} \dot{x}_1 &= 5x_2 + 5x_3 - 16x_1 + u_1 \\ \dot{x}_2 &= 14x_1 - 95x_2 \\ \dot{x}_3 &= 90x_2 - 14x_3 \end{aligned} \tag{5}$$

were generated to a finite arithemetic precision (d = 5) with uniform white noise added to the last digit in order to simulate experimental error. The model and statistical details of experiment are identical to analysis performed by Mulholland and Weidner (1980) for the regression analysis identification problem. This allows for a comparison of the results obtained with Prony's method with those of regression analysis. For more details concerning the Monte Carlo experiment the reader is referred to the paper by Mulholland and Weidner (1980).

Using the noisy samples of (5), Prony's method was employed to compute the a_{11} and a_{12} matrix entries defined by (5) for approximately 20 different values of the sample period τ between 0.002 and 0.12. For each sample period

value, 400 simulations were used to compute the model parameters. The sample means and variances for a_{11} and a_{12}, two particularly sensitive parameters, were computed for each value of τ. Figures 1 through 4 are the plots of these sample moments versus τ. Note that in Figs. 1 and 3 the sample means all lie within 10 percent of their true values, $a_{11} = -16$ and $a_{12} = 5$, regardless of the sample period. However, the sample variances, plotted in Figs. 2 and 4, show a concave functional dependence on τ.

The plots of the sample variance clearly illustrate the existence of a window of acceptable sample periods for model identification. A very similar result was obtained by Mulholland and Weidner (1980) for model identification using regression analysis. For Prony's method the sampling window is shifted slightly to lower values of τ relative to the window obtained with regression analysis. Prony's method requires a longer time record for model identification, hence shorter sampling times should be observed for the same sample variance values (Mulholland and Weidner, 1981).

Flow-Through Study

In March 1979 at the Environmental Research Laboratory (Athens, GA) a flow-through microcosm experiment was begun, in which simultaneous step and impulse inputs of the pesticide methyl parathion were applied. Measurements of only the radiolabeled pesticide and its products in water were made before and after the time at which the inputs were applied. In all six microcosms (tanks) were exposed to the pesticide. Figure 5 gives a sample response for the pesticide in the water compartment of tank 2.

Figure 5 gives approximately 35 days of the time series of measurements for the microcosm experiment. The data collected extends for 42 days beyond that shown, however these measurements merely establish the random nature of the steady-state entrained by the constant infusion of radiolabeled pesticide.

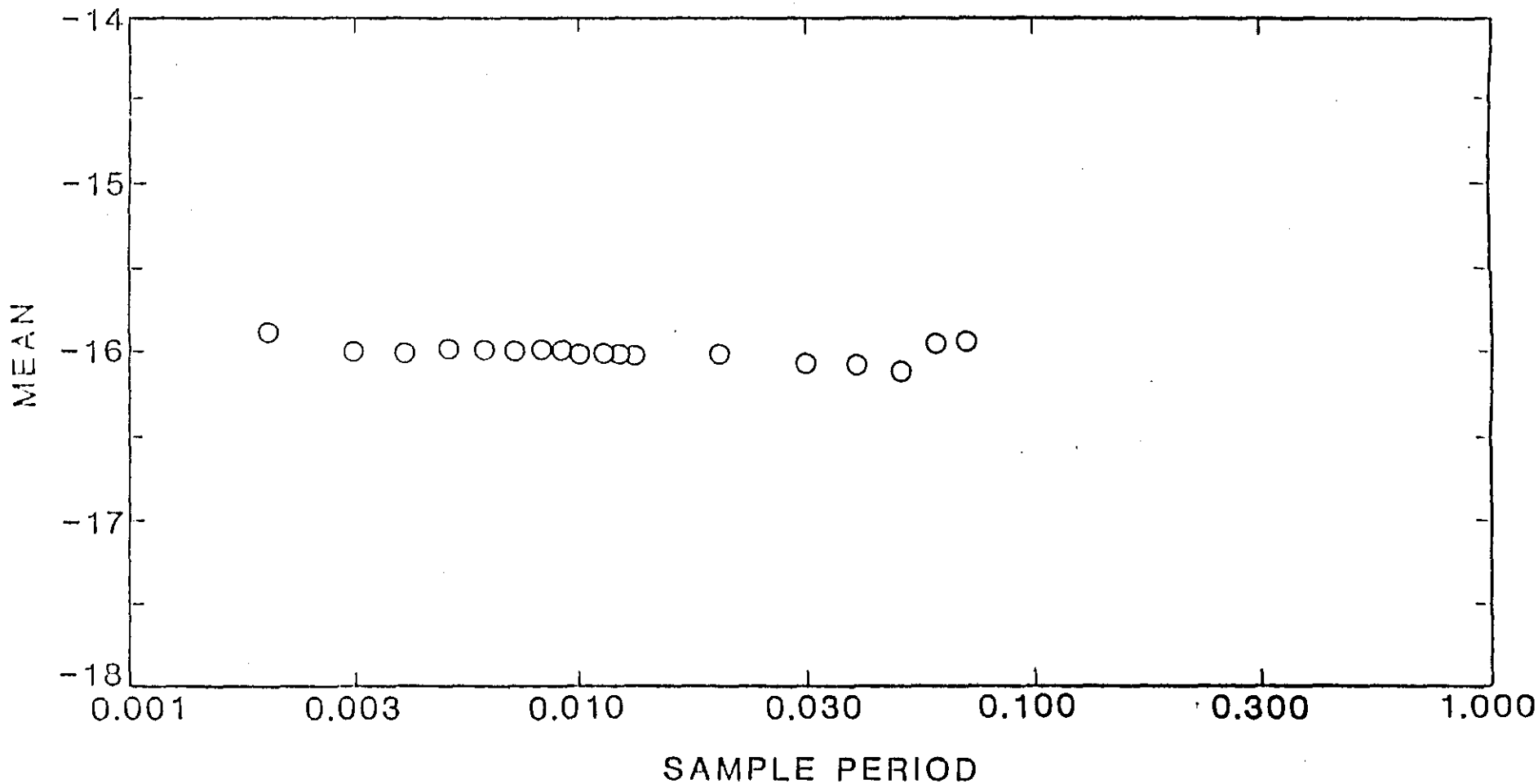


Figure 1. Estimate for the mean value of parameter a_{11}.

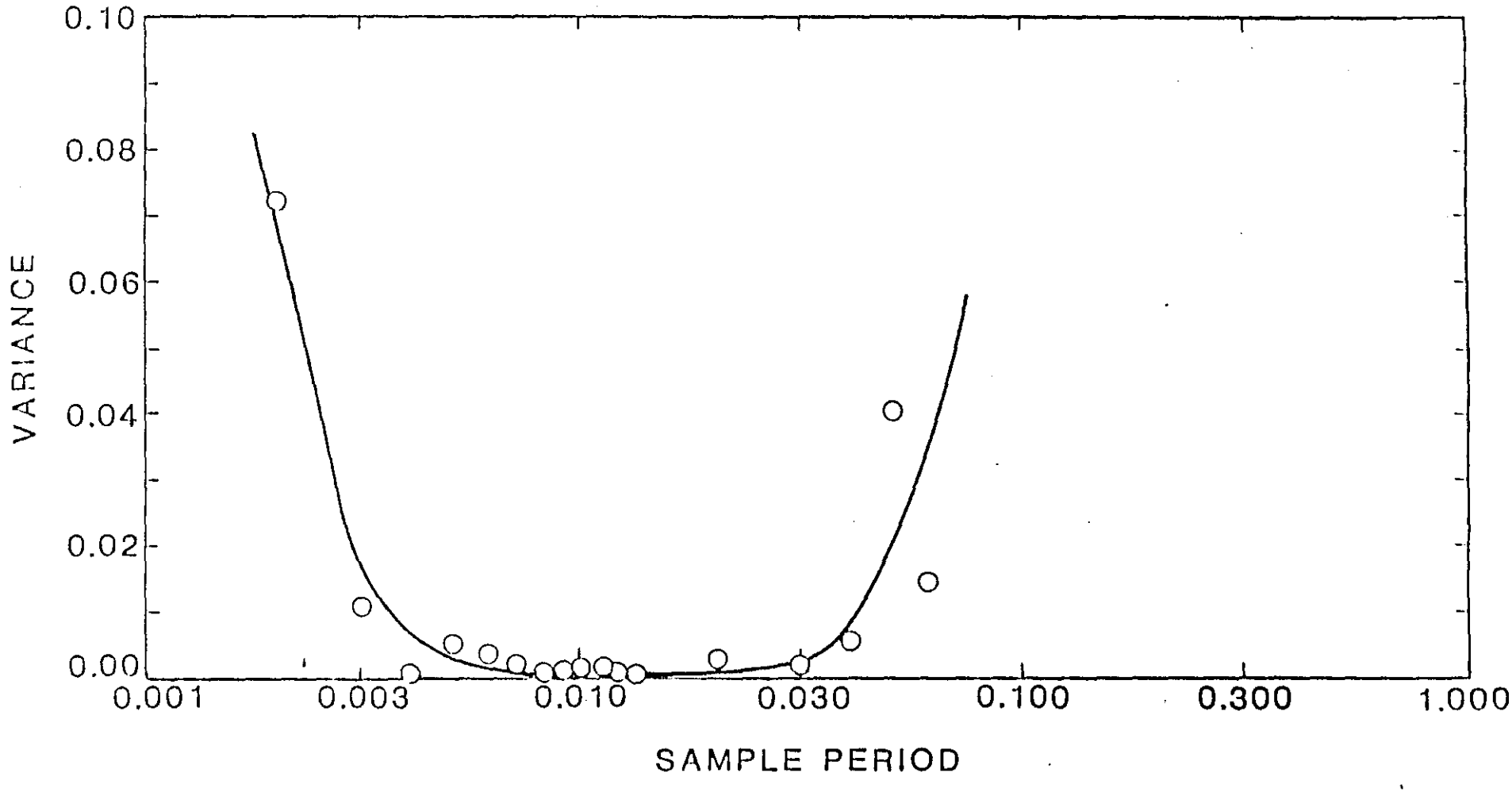


Figure 2. Estimate for the variance of parameter a_{11}.

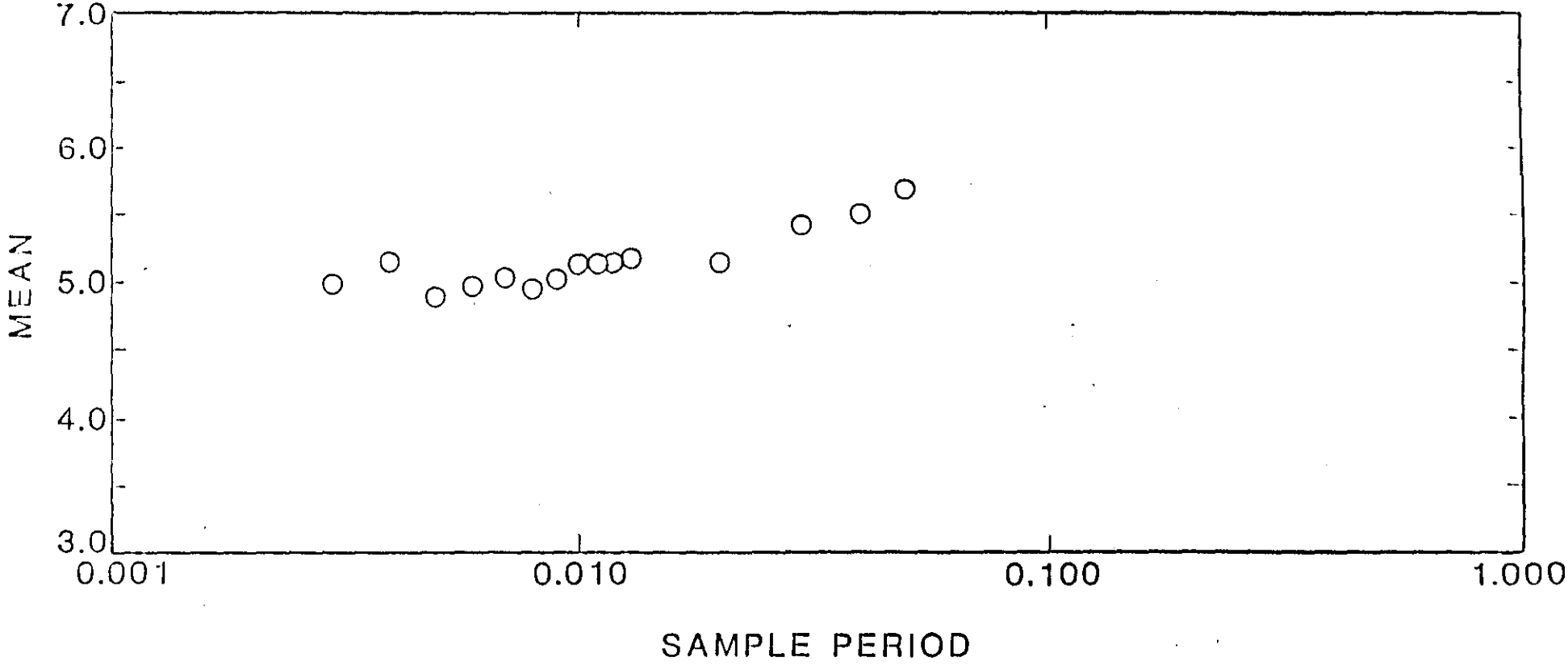


Figure 3. Estimate for the mean value of parameter a_{12}.

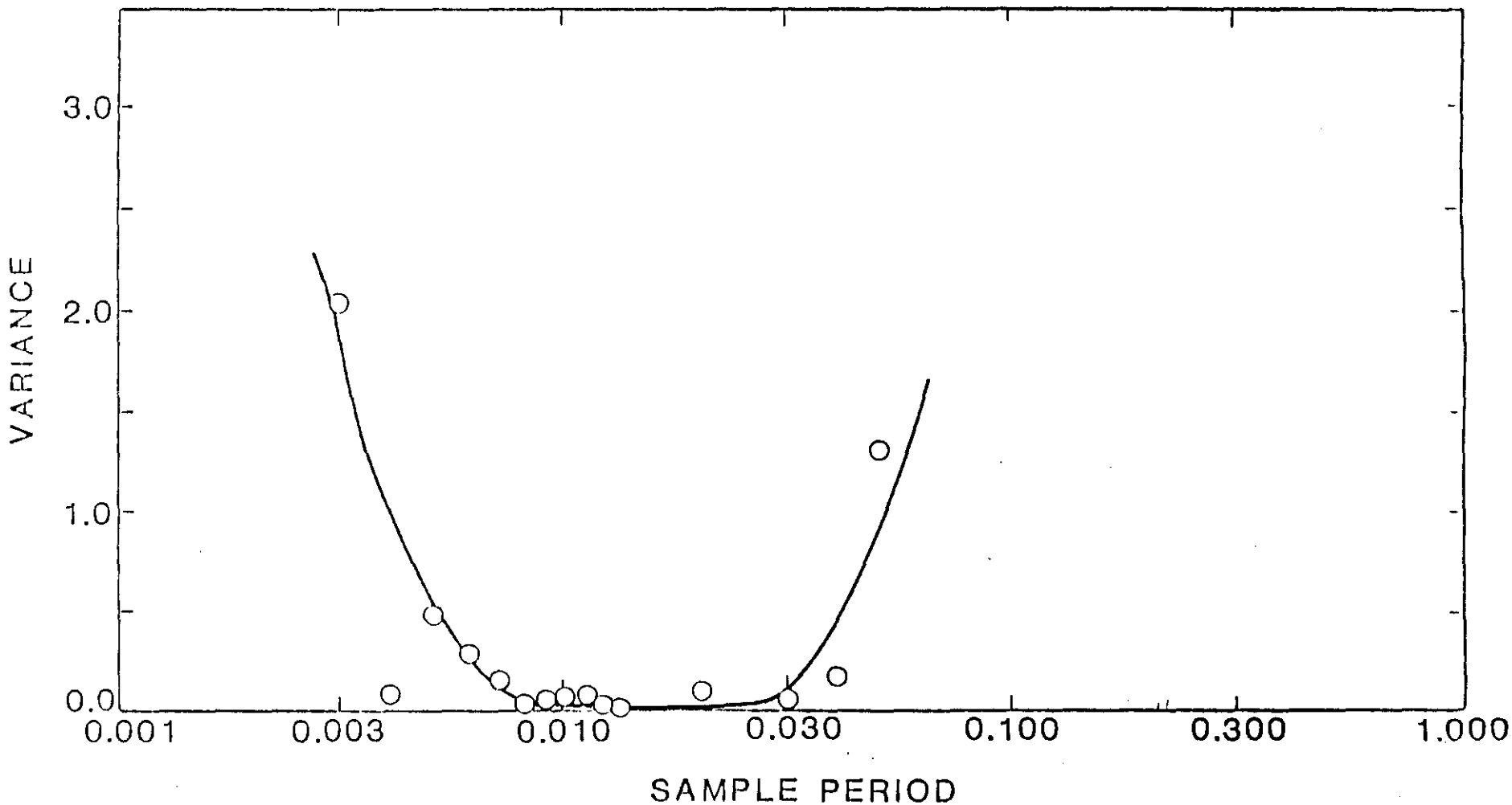


Figure 4. Estimate for the variance of parameter a_{12}.

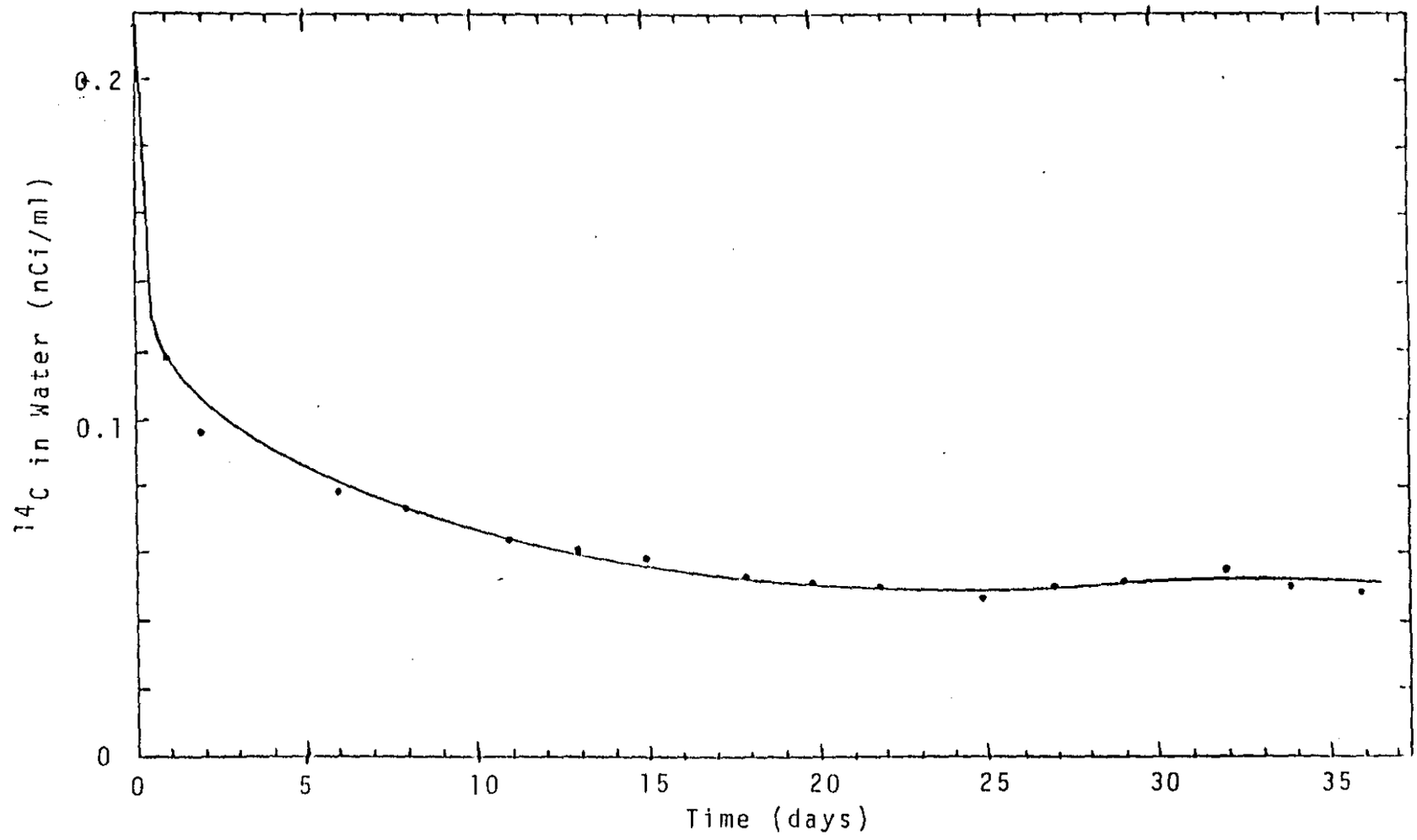


Figure 5. Water compartment response (tank 2).

The data set also includes measurements made for several days before the inputs were applied. These data establish the measuring instrument error, which never exceeds 0.002 nCi/ml.

The data for the first six tanks is the measured response of the water compartment to the simultaneous application of single dose and constant inputs. The response was determined to be of the form

$$x_1(t) = p_0 + p_1 e^{\lambda_1 t} + p_2 e^{\lambda_2 t} \qquad (6)$$

where the coefficients p_0, p_1, and p_2; and the eigenvalues λ_1 and λ_2 are real constants. For tank 2 (see Figure 5) it is clear that the constant term p_0 dominates after 18 days. This term is identified by taking the average value of the data for $t \geq 18$ days. This value is then subtracted from the data for tank 2 leaving the two exponential terms to be identified from the modified data.

Now assume that $|\lambda_1| > |\lambda_2|$, and call λ_1 the fast eigenvalue and λ_2 the slow eigenvalue. It is clear that after two days the λ_1-term is essentially zero within the precision of the measured data. This means the sample period of one day is not sufficiently short to capture the dynamics of the microcosm for Prony's method. Verification of this observation has been obtained by a direct application of Prony's method, which fails to produce a real number for λ_1. Another (graphical) technique, called curve peeling, was then applied to the data.

Between 2 and 18 days the logarithm of (6) should be a linear function of p_2 and λ_2. These two constants were computed by fitting a linear curve to the data. And, the constants p_1 and λ_1 were obtained by fitting the fast eigenvalue mode to the initial (first two days) data. The results of these data fits give the parameters of $x_1(t)$, defined in (6), in Table 2.

The results of the curve peeling provide valuable insight into the failure of Prony's method. Analysis by Prony's method requires a minimum of

TABLE 2. MICROCOSM CURVE PEELING RESULTS

Tank	λ_1	λ_2	P_o	P_1	P_2
1	-1.33	-0.072	0.048	0.113	0.042
2	-3.28	-0.177	0.054	0.080	0.074
3	-1.04	-0.086	0.054	0.119	0.029
4	-2.60	-0.133	0.057	0.081	0.065
5	-1.89	-0.315	0.045	0.053	0.083
6	-2.45	-0.193	0.044	0.081	0.080
average	-2.10	-0.163	0.050	0.088	0.062
standard deviation	0.840	0.089	0.0054	0.244	0.0219

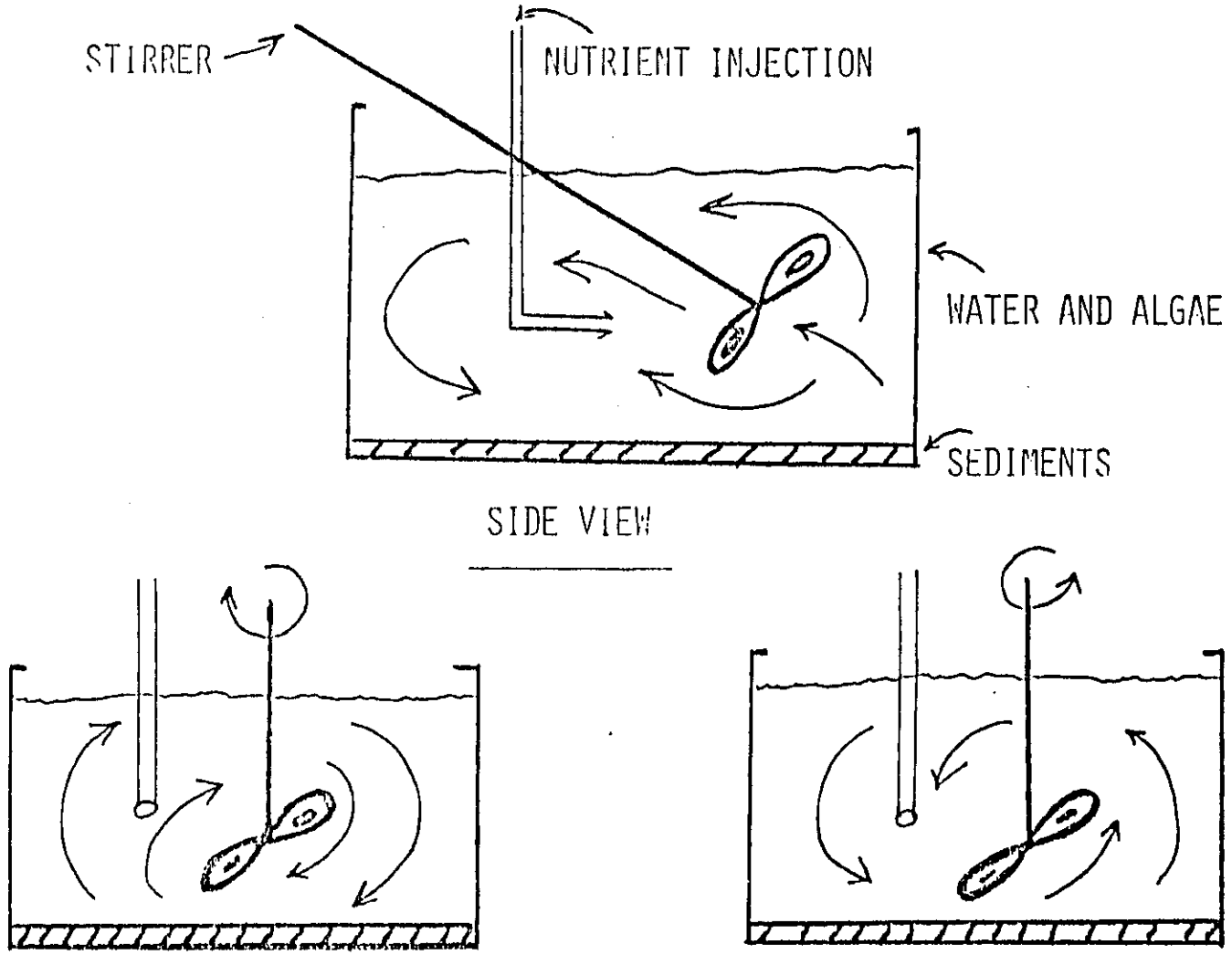
2n time periodic data points where n is the number of exponential terms to be identified. For (6), four points are required. Suppose that measurements for the initial condition and for each of the first three days are used. From the curve peeling technique, the contribution of the term $p_1 e^{\lambda_1 t}$ on the third day is

$$0.080^{-3.28(3)} = 0.0000043$$

for tank 2. Since the measured data is limited to 3 decimal places, the contribution of the fast exponential does not appear in the data for the third day, nor does it appear on the second day. Thus, Prony's method, in this case, requires the use of data which contains no information regarding one of the terms to be identified. When applied to the data for tanks 2 and 4, Prony's method gives one real eigenvalue and one complex eigenvalue, an erroneous result for real systems. Clearly, the sampling period is too long for the precision of the measurements. Selection of a proper sampling period for Prony's method is apparently limited by a maximum value of one-half day.

Finally, something very unusual was discovered in the curve peeling results presented in Table 2. A simple calculation produced the result that a goodness-of-fit of the models for the even-numbered microcosms was much better than for the odd-numbered microcosms (61% vs. 93%). Fisher's F-test on the odd and even pooled mean-square-errors showed the variances were significantly different ($p = 0.005$). Subsequently, it was noted that stirrer rotation was counter-clockwise (CCW) for odd and clockwise (CW) for even microcosms (see Figure 6). There were no other discernable differences.

Since the pitch of propellers were opposite for CW and CCW rotation, the mixing in odd and even were mirror images. However, the nozzles for injecting nutrient media and toxicant were all on the left sides of the stirrer shafts. Thus, nutrient input in CW units were entrained in a vortex that interacted with wall growth and algal mat first and in CCW units it



AQUATIC MICROCOSMS AT ATHENS EPA LABORATORY

Figure 6

interacted with sediments first. Three observations from ancilliary experiments combine to explain the difference between the variance of system-level properties in CW and CCW stirred microcosms.

Experiments with an algal community subsampled from the microcosms indicated that measures of functional properties (e.g., toxicant degradation rate, nutrient uptake, growth curve) had low variability within replicates (c.v. = 10%). Analogous experiments with sediments showed a greater variability (c.v. = 60%). A set of experiments on partitioning of organic compounds between silicone rubber and microcosm sediments suggests a two-stage sorption process. The first stage is rapid ($t_{\frac{1}{2}} \approx 2$ hr) and clearly reversible. The second stage is shower ($t_{\frac{1}{2}} \approx 2$ days) and "irreversible" on our time scale of 21 days.

The interaction of incoming nutrients and toxicant with functionally homogeneous algal communities in the even microcosms and the interaction of incoming nutrients with functionally heterogeneous sediments with a relatively irreversible process are sufficient to explain the difference between variances of system-level properties in the CW and CCW stirred microcosms. Additional experiments at the Athens EPA Lab are being designed to test this hypothesis.

The subtle relationship between stirrer rotation and input location changed the mean-square-error deviation from the identified model by over 400%, and variance of other parameters changed comparably (300 to 500%). The microcosm apparatus has been modified to eliminate this source of variability. This and other changes over the past 8 years are primarily responsible for improvement in explanatory power of the models from less than 20% to more than 90% of the variability in the Data. From this it is observed that highly replicable and repeatable experiments with microcosms as surrogate

ecosystems demand careful, competent, experimental development and methods. Without prior experience, a one or two year project to develop a meaningful data set from microcosm experiments should be suspect.

While the variance of model coefficients and system parameters increased by 300 to 500% in the CCW microcosms over the CW microcosms ($p = 0.005$), the means of the same coefficients and parameters were indistinguishable ($p = 0.01$). Thus, the theoretical assertion that the mean and variance of a variable are independent is correct for system-level properties of our microcosms. The following conclusion is obvious: Analysis of response to a perturbation of the variance of an input to the microcosms may be performed independently of the mean of the input which sustains nominal ecosystem behavior.

V. CONCLUSIONS AND RECOMMENDATIONS

The results of this research have produced an experimental design for the laboratory analysis and evaluation of perturbed microcosms. The complete design has not yet been fully tested against actual experiments. A continuation of this line of research would allow such a test to take place. Additional research is also needed for the further development of a stochastic version of the experimental design, thus combining statistics with systems analysis, which could lead to a powerful technique for error analysis. The development of a statistical error analysis should give rise to techniques for filtering raw data, and hence for minimizing measurement errors in the presence of noise. Testing of mathematical models against microcosm derived data, based upon the proposed development of a stochastic experimental design, could lead to a rigorous methodology for model validation.

Preliminary data obtained from experiments conducted at the Environmental Research Laboratory, Athens, Georgia have been analyzed. These data measure the effects of pesticides (for example methyl parathion) on the

structure of aquatic microcosms. The rate coefficients of the water compartment of a laboratory ecosystem receiving toxic inputs of methyl parathion have been computed. This research has disclosed a method for computing bounds on the time required for a pollutant to flush from a system after the source of pollution has been removed. The bounds depend on the system turnovers and the minimum level of pollutant detection using prescribed measurement techniques. Work in progress seeks to find new ways to compute the maximum acute toxicity resulting from an introduced slug of toxic material passing through an aquatic system.

Applications of the results of the research project clearly focus upon several water problems of Oklahoma as described in 5-year plan for water research of the Oklahoma Water Resources Research Institute. The research describes techniques which will aid in characterizing the toxicological effect of effluents which enter both surface and ground water systems. In particular, runoff from pesticide-treated non-point source agricultural areas is a water pollution problem which this research program addresses.

Finally, this project has described a first step in the use of systems analysis and modeling for the design of laboratory experiments. The use of statistics when combined with systems theory can be a powerful tool as was demonstrated in the analysis of the microcosm curve peeling results. It now remains to transfer these techniques to the analysis of water pollution problems in natural systems.

ACKNOWLEDGEMENT

The author wishes to acknowledge his friendship and long term scientific collaboration with Dr. James Hill, IV of the US EPA Environmental Research Laboratory in Athens, GA. Many of the ideas presented in this report resulted from discussions with Dr. Hill.

REFERENCES

Hildebrand, F. B., 1956. Introduction to Numerical Analysis. New York: McGraw-Hill.

Muir, K. T., 1980. Nonlinear Least-Squares Regression Analysis in Pharmacokinetics, Comp. Biomed. Res. 13:307-316.

*Mulholland, R. J., J. L. Crittenden, and R. J. Weidner, 1980. Effect of Data Sample Period on Model Identification. Proc. 1980 Int. Conf. Cybernetics and Society, October 8-10, 1980, Boston, MA.

*Mulholland R. J., 1981. New Sampling Theory for Measuring Ecosystem Structure, OSU Center for Systems Science Report 82-2, Stillwater, OK.

*Mulholland, R. J. and R. J. Weidner, 1980. Sampling Strategies for the Identification of Compartment Models, Int. Jour. Systems Science 11:377-387.

*Mulholland, R. J. and R. J. Weidner, 1981. Stochastic Properties of Compartment Models, 12:927-936.

Prony, R. 1795. Essai Expérimental et Analytique sur less Lois de la Force Expansive de la Vapeur de l'Alkool, á Differentes tes Températures. J. L'Ecole Polytech. 1:24.

*Papers produced as a result of this grant.