SETAC PRESS

Environmental Toxicology and Chemistry, Vol. 20, No. 9, pp. 1942–1950, 2001

Printed in the USA
0730-7268/01 $9.00 + .00

PERSISTENCE OF THE EFFECTS OF JET-A IN A MICROCOSM WITH RELEASES FROM THE SEDIMENT

RANDY S. SANDBERG and WAYNE G. LANDIS*
Institute of Environmental Toxicology and Chemistry, Huxley College of Environmental Studies, MS 9180, Western Washington University, Bellingham, Washington 98225-9180, USA

(*Received* 7 *April* 2000; *Accepted* 16 *January* 2001)

Abstract—This study investigates both the methods and the multispecies and functional level effects of a release of a complex hydrocarbon mixture from sediments using a 60-d modified mixed flask culture (MFC) microcosm. Neat Jet-A was injected and mixed into the sediment with one nondosed and three dosed concentrations. Univariate and multivariate statistical and graphical techniques were used to detect patterns in the data. A slow release of the test material from the spiked sediment layer was obtained, and constituents of Jet-A were detected. Functional parameters (such as pH) were generally better at determining treatment groups than structural parameters (population densities). Analysis of the ability of the various parameters to detect treatment differences confirms that there is not one best indicator for the status of an ecological structure. Transient but statistically significant outcomes were seen at initial treatment concentrations as low as 2 μl/L. The higher concentration treatment groups could be identified as distinguished from nondosed or lower treatments at the end of the 63-d experiment. Each of the three multivariate techniques differed in their ability to distinguish treatment groups during the course of the experiment.

Keywords—Jet-A Sediment Mixed flask culture microcosm Multivariate analysis Community conditioning

INTRODUCTION

This report details the persistent effects of neat Jet-A introduced to the sediment of a mixed flask culture (MFC) microcosm experiment. Our results demonstrate persistent and varied effects on functional and structural components due to the toxicant. Effects were seen in the microcosm experiment at concentrations 100 times less than the amphipod median effective concentration (EC50) sediment toxicity test using the same type sediment and media.

Sediments are a major repository for contaminants introduced into surface waters [1]. We now recognize that physiochemical and biological relationships between sediment contamination and the sediment environment are complicated and not easily managed using chemical criteria. This has led to increased monitoring of sediment contamination and benthic macroinvertebrate communities by regulatory agencies [2,3].

Burton [4] criticizes sediment toxicology due to the failure to incorporate ecosystem disturbance into toxicity assessments. Numerous single-species assays have been developed for the assessment of sediment toxicity to a variety of organisms. Although this testing satisfies the objectives of defining sample toxicity to the test species, these tests do little to document and define toxicity to components of ecological systems. Significant cases of acute toxicity are encountered infrequently [5], while subacute levels of contamination with the potential to disrupt ecosystem structure and function are more common.

Researchers criticize the approach of relying on single-species tests because these tests may not be adequate predictors of potential effects on communities and ecosystems [6,7]. Within a laboratory test, acute lethality values vary with species, strain, age, environmental conditions, genomic structure, and other confounding factors. Many more species are present in an ecological system than are used in single-species toxicity testing. Extrapolation is required typically using species concentration–response regressions with an associated uncertainty. Even if it were possible to test all the individual species, single-species toxicity tests do not represent ecological systems.

The interactions that define an ecological system are not found in single-species toxicity tests or bioassays. The potential rate of increase for a field population, *r*, will decrease not only with an increase in the death rate predicted by the EC20 or median lethal concentration (LC20) but also a shrinking birth rate through interference in fertility, fecundity, and development due to indirect effects. Indirect effects are not present in single-species toxicity tests. Organisms within the ecological systems can degrade, biotransform, or accumulate the xenobiotic, thereby altering the exposure pathway. Physical components such as sediment or suspended organics can bind to the toxicant. In a nutrient-rich state, the resident organisms may overcome the toxic effects of a xenobiotic. Population blooms can occur as grazers, and predators are killed. To evaluate these interactions, we must conduct field research or use an experimental model incorporating the interactions of interest. The experimental models are multispecies toxicity tests, typically called microcosms and mesocosms.

Over the past 30 years, researchers have developed a variety of multispecies aquatic toxicity tests that incorporate a subset of the features found in the field [8–12]. The size of multispecies tests can range from 1-L microcosms, as in the MFC [9], to the thousands of liters commonly used in pond mesocosms for pesticide registration testing. Multispecies toxicity tests provide the interactions missing from single-species toxicity tests and the replicability not possible in field studies. Problems exist with the data analysis of both field studies and multispecies toxicity tests.

One major difficulty in the evaluation of multispecies tox-

* To whom correspondence may be addressed (landis@cc.wwu.edu).

icity tests is the analysis of a large multivariate data set with relatively few replicates compared to the number of features. Microcosm studies of the 1990s spurred development of a variety of univariate and multivariate statistical and graphical tools [13–20].

Analysis of variance (ANOVA) is a classical method to examine single-variable differences from the control group. However, because of the temporal nature of multispecies testing, problems exist with using conventional ANOVA [20–23]. These include the increasing likelihood of introducing a type II error (accepting a false null hypothesis), temporal dependence of the variables, and the difficulty of graphically representing the data set. Conquest and Taub [21] developed a method to overcome some of the problems by using intervals of nonsignificant difference (IND). The IND produces upper and lower limits of a variable for each sample date that demonstrate the boundaries beyond which the treatment was statistically significant as determined by ANOVA and a multiple comparisons test. This method corrects for the likelihood of type II errors and produces intervals that are easily graphed for ease of analysis. The resulting graphs portray when the variables in the dosed treatments are statistically different from in the nondosed treatments. The major drawback of this technique is the need to examine multiple single variables over the course of the experiment, thereby relying heavily on the expertise of the evaluators in determination of the final result.

Multivariate methods are especially useful in ascertaining patterns of effects in microcosm experiments. Landis et al. [19,22,23] used metric and nonmetric clustering and association analysis (NCAA) as developed by Matthews, Matthews, and coworkers [24–26]. Other multivariate techniques have been developed and used extensively in the evaluation of stream microcosms [12,27].

Graphical tools allowing the visualization of dynamics are one of the best methods of data analysis. Classical plots of species density over time are useful especially in concert with the graphical INDs [21]. Area graphs depicting the species composition of certain components of the experimental system can provide the means to interpret changing community structure [22,23]. Plots of dynamics in phase space, plotting the density of one component by another, can allow the visualization of patterns of interaction that may be missed using conventional presentations [28]. Landis et al. [18,29] used three-dimensional representations of these phase plots over time.

Kersting and Van den Brink [27] have used an alternative method to present the results from a series of microcosm experiments. Redundancy analysis can be used to construct a two-dimensional representation of the multivariate space describing the ecological structure of the microcosm. The position of the treatments can be displayed on the graph, and the trajectories, including convergences and divergences, can be observed. A sense of the relative dynamics of the treatment groups can be determined.

The classic criticism of multivariate toxicity tests and ecological microcosm experiments is that they do not model all the characteristics of a field ecological structure [30–32]. This criticism can be applied to all models, including single-species toxicity tests, microcosms, simulation models, or enclosure experiments performed in the field. In each case, a simplification allows us to improve tractability, to increase statistical power, and often to test a specific hypothesis where the spatial and temporal structure of the field is not necessary. Field testing of pesticides and other contaminants at relevant ecological scales has obvious drawbacks.

The relative sensitivity of functional variables versus structural measures in microcosm systems is also of interest. Functional variables are those measures that describe the metabolism of an ecological system. Dissolved oxygen dynamics, nutrient concentrations, pH, and absorbance are classical examples. Structural variables are those that describe the dynamics of the organisms within the microcosm. Species densities, relative abundance of organisms, and diversity measures are typical examples. Kersting and Van Wijngaarden [12] and Van Geest et al. [33] discuss the relative importance of these measures in a series of herbicide mesocosm experiments. Landis et al. [19], in an experiment with the turbine fuel JP-8, found that pH was often an important variable in determining treatment effects but that structural components were also critical.

This paper describes the use of a natural community incorporated in an MFC to evaluate the effects of a release of Jet-A turbine fuel from sediments using both conventional univariate and multivariate analyses. Analysis of the Jet-A during the course of the test revealed a slow release of the test material from the spiked layer. Community functional parameters revealed an initial treatment effect apparently caused by both the transfer perturbation of the spiking procedure and the toxic effects of the hydrocarbon mixture. Treatment effects were generally detectable throughout the entire test with no apparent recovery, as defined as returning to the state of the nondosed replicates. Multivariate techniques were able to distinguish treatment effects during most of the experiment. Several of the structural and functional parameters showed concentration-related effects at the end of the experiment.

MATERIALS AND METHODS

Two sets of toxicity tests were conducted. A short-term sediment toxicity test was performed in order to set concentrations for the MFC experiments and for comparison. The same amount of sediment and media were used in each experimental replicate in both the short-term toxicity test and the microcosm experiments. The concentrations reported are for the total volume of the experimental replicate, 1 L in each case.

Acute tests

We conducted a series of short-term sediment toxicity tests using *Hyalella azteca* as specified by ASTM E 1383-90 [34] to set MFC concentrations and for comparison. We used the same sterile sediment, media, and species for the toxicity tests as used in the MFC.

Individual test chambers were spiked with Jet-A added to 100 ml of silica sand sediment by injection using a Hamilton chromatography syringe (Hamilton Supplies, Reno, NV, USA). Immediately following this, the sediment was stirred with a sterile glass rod, the chamber covered with a 150 × 15-mm-diameter petri dish to minimize evaporation of the hydrocarbons, and homogenized for 15 s with a vortex mixer. To avoid mixing during filling with sterile T82MV media [35], a 100 × 15-mm-diameter sterile petri dish was placed in the chamber over the sediment. The petri dish was then carefully removed with sterile forceps to minimize disturbance of the sediment–water interface. Testing with *H. azteca* was then carried out using the standardized protocol [34]. The definitive toxicity test was for 10 d with six replicates with four treatments of

0, 250, 500, and 750 μl of Jet-A. The total volume of media and sediment was 1 L. Mortality was the endpoint. The resulting data were subjected to an appropriate definitive statistical analysis under a null hypothesis of treatment having no effect. The LC50s were determined graphically and by probit analysis when appropriate.

Mixed flask culture

Leffler [9] describes the MFC protocol. Briefly, test chambers containing 50 ml of SAM sediment, 900 ml of T82MV sterile media, and 15 μg/L of $NaHCO_3$ were inoculated with 50 ml of a naturally derived stock community and kept at 20 ± 1°C on a 12:12-h light:dark schedule for six weeks. The stock community was established via inoculations from freshwater ponds and streams in western Whatcom County (Washington, USA). To improve similarity among replicates, we conducted cross-inoculations once a week and rotations in the incubator twice weekly during this period. The cross-inoculations ceased at the dosing of the microcosms.

On day 0 of the experimental period, individual test chambers were spiked with Jet-A according to treatment group in a manner similar to the acute tests. New and identically cleaned and numbered 1-L chambers, containing an additional 50 ml of standardized aquatic microcosm (SAM) sediment, were injected with Jet-A using a Hamilton chromatography syringe. The 0-μl or nondosed group received only distilled deionized water. Treatments 2, 3, and 4 received 2.0, 10.0, and 25.0 μl of Jet-A, respectively, with six replicates per treatment. After mixing and vortexing, a petri dish was immediately placed in the new chamber over the treated sediment using sterile forceps. The covering of the sediment by the petri dish prevents mixing of neat material on adding the conditioned microcosm and sediment. The original microcosm was transferred over to the new dosed chamber by gently pouring and scraping with a sterile rubber policeman. The petri dish was gently removed using sterile forceps to minimize disturbance of the underlying spiked sediment.

Sampling was then carried out in accordance with the established SAM [35] and MFC [9] protocols. Sampling included dissolved oxygen, pH, turbidity, and organism numerical densities twice weekly. Dissolved oxygen was monitored as specified in the SAM protocol to calculate photosynthesis-to-respiration (P:R) ratios. Numerical densities of the biota in each chamber were determined by subsampling for algal, protozoan, and large organism counts utilizing the devices and procedures specified by the SAM protocol. The sampling process for the organisms resuspends most of the sediment into the water column.

Gas chromatography of the Jet-A components

Two samples from each treatment group were also collected each sampling day for use in tracking the pulsing in concentration of the Jet-A in the dosed treatment groups as a result of the disturbance of sampling. Four milliliters of media were removed from the approximate center of each sampled chamber using a 10-ml disposable pipette and stored at 4°C in a cleaned and acid-washed screw-top test tube. These samples were subsequently analyzed using purge and trap (P&T) gas chromatography. This was performed using a Tekmar LSC 2000 P&T concentrator (Cincinnati, OH, USA) in tandem with a Hewlett-Packard 5890A gas chromatograph (GC) and a flame ionization detector (Avondale, PA, USA). Deionized distilled water blanks were used to verify the P&T and GC columns' cleanliness prior to analysis of the sample. A 3.5-ml sample was injected into a 5-ml sparger, purged with prepurified nitrogen gas for 11 min, and dry purged for 4 min. Volatile hydrocarbons, purged from the sample and collected on the Tenax/Silica Gel column (Tekma-Dohrman, Cincinnati, OH, USA), were desorbed at 180°C directly onto the gas chromatograph SPB-5, 30-m × 0.53-mm i.d. 1.5-μm film, fused silica capillary column. The column at 35°C was held at that temperature for 2 min, increased to 225°C at 12°C/min, and held at that temperature for 5 min. A Spectra-Physics 4290 Integrator (Mountain View, CA, USA) was used to record the flame ionization detector signal output of the volatile hydrocarbons that were separated and eluted from the column by molecular weight.

Microcosm data analysis

Data from the MFC were recorded on computer entry forms and then entered into a computer. Entries were checked for accuracy, and numerical densities of each monitored biotic category were calculated along with net photosynthesis (P), respiration (R), photosynthesis/respiration ratio (P:R), absorbance (A), and total algae in accordance with the SAM protocol [35]. Total protozoa and total invertebrates were calculated in a manner similar to total algae.

The univariate statistical significance of all calculated parameters, along with the physical parameter data for dissolved oxygen and pH, were computed using the IND as developed by Conquest [21]. As specified in the SAM protocol, ANOVAs were calculated each sampling day for each variable and were used to plot average daily values and INDs over time to identify significant differences between the controls and treatments under a null hypothesis of treatment having no effect.

The goal of the multivariate methods was to test if the microcosm's replicates could be assigned to a treatment group looking at a suite of characteristics. Three previously described multivariate analysis techniques used microcosm data analysis [20,22–24,26]. Two of these methods were based on the ratio of multivariate metric distances within treatment groups versus treatment groups using the distance measures of cosine vector (an angular measure) and Euclidean distance between test chambers [36,37]. Statistical significance of the clustering using these two procedures was determined by analyzing the average within- and between-group distances using a permutation test [38]. This test is based on the fact that if the treatment has no effect, assignment of points to treatment groups will be random. The procedure randomly assigns each of the replicate points to groups and recomputes the within-group to between-group ratio (W:B) many times. If the null hypothesis is false, this randomly derived ratio will (probably) be larger than the W:B ratio obtained from the actual treatment groups. By taking a large number of random reassignments, an estimate of the probability under the null hypothesis is obtained as $(n + 1)/(500 + 1)$, where n is the number of times a ratio less than or equal to the actual ratio was obtained [38].

The third method, RIFFLE, utilizes a NCAA algorithm [39]. The nonmetric clustering and the RIFFLE program have been used with success by Landis and colleagues in the analysis of a variety of microcosm data sets [22,23]. Nonmetric clustering uses the characteristics of the replicates to assign them to clusters and is naive of treatment group. The characteristics important in assigning the replicates to clusters is documented. In order to test if the clustering is related to treatment, a simple observed–expected contingency χ^2 good-

Table 1. Parameters measured and calculated in the mixed flask culture (MFC). The parameters used in each of the clustering methods are marked. The totals for each group, algae, protozoa,and invertebrates were not used in the clustering since they are composites of other variables

	Metric clustering	NCAA[a]
Functional parameter		
pH	x	x
AMDO1 (first morning before lights on dissolved oxygen)	x	x
PMDO2 (afternoon dissolved oxygen)	x	x
AMDO3 (second morning before lights on dissolved oxygen)	x	x
Absorbance photosynthesis-to-respiration ratio = (AMDO1 − PMDO2)/(PMDO2 − AMDO3)	x	x
Structural parameter		
Selenastrum sp.	x	x
Chlorella sp.	x	x
Scenedesmus sp.	x	x
Ankistrodesmus sp.	x	x
Other green algae	x	x
Filamentous green	x	
Nitzchia sp.	x	x
Other diatoms	x	x
Lyngbya sp.	x	x
Other bluegreens	x	x
Total algae		
Amoeba	x	x
Ciliates	x	x
Flagellates	x	x
Paramecium bursaria	x	x
Total protozoa		
Rotifers	x	x
H. azteca	x	x
Copepods	x	x
Ostracod 1	x	x
Ostracod 2	x	x
Insect larvae	x	
Total invertebrates		
Bacteria	x	

[a] NCAA = nonmetric clustering and association analysis.

ness-of-fit association analysis is used to determine the significance of the clustering produced by RIFFLE.

Because of the suspected dependence of the community functional variables on structural variables, multivariate analysis was performed on functional and structural components separately. The derived variables of P, R, P:R ratio, total algae, total protozoa, and total invertebrates were excluded because of their dependence on other included variables. Table 1 lists the parameters used for the multivariate analyses.

Significance levels generated from these procedures were plotted over time in order to examine the presence of treatment-related effects on each sampling date. The comparison of the three methods allows more accurate determination of the presence and persistence of the formation of clusters related to treatment effects.

RESULTS

Acute tests

We calculated the toxicity of the Jet-A spiked sediments on a volume-per-volume (v/v) basis to be comparable to the microcosm protocol. Graphically obtained LC50s for the preliminary and definitive tests were approximately 512 and 263 μl/L of Jet-A. A probit analysis of the definitive test yielded an LC50 of 259 μl/L. The LC20 was estimated graphically at 125 μl/L. Water quality characteristics for each test were well within the limits of acceptability as defined in the acute toxicity test protocol.

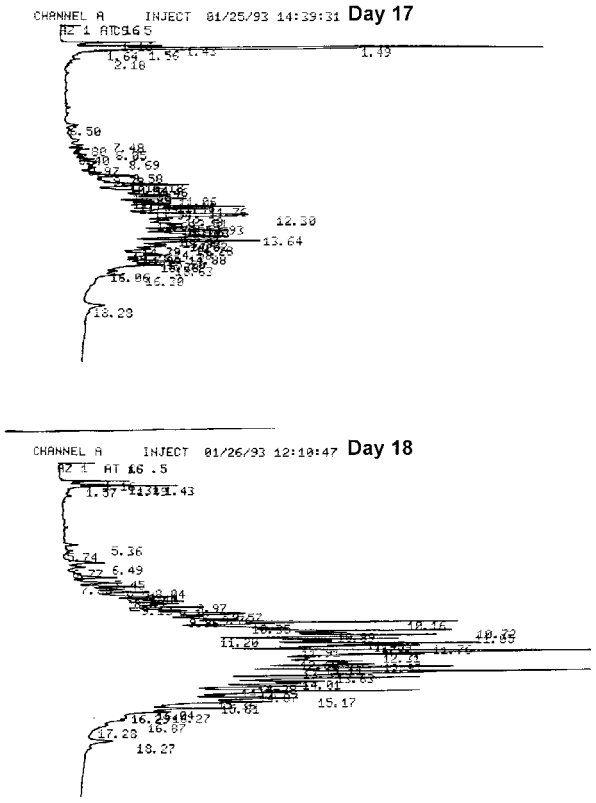

Fig. 1. Gas chromatography results from the 25-μl treatment group on days 17 (January 25, 1993) and 18 (January 26, 1993). Day 17 was taken prior to sampling on day 18. The increase in concentrations of the Jet-A components is apparent after the stirring during the sampling of the organisms.

Jet-A measurements

Results obtained from the GC analysis of the Jet-A components indicated that stirring the sediment during sampling released the Jet-A and essentially redosed the spiked groups during sampling periods (Fig. 1). Although no clear exposure duration could be determined, the Jet-A remained in the test systems for a substantial portion of duration of the test. Results from both the 10- and 25-μl treatment groups on day 25 indicated that a portion of the turbine fuel was still in the test systems.

Effects in the MFC

At day 0, the replicates randomly assigned to treatment groups were similar enough that no treatment-related pattern was apparent by either univariate or multivariate techniques. Variability as judged by the IND was also low. After the cessation of cross-inoculation and dosing with Jet-A, an increase in variability, treatment-related effects, and clustering was seen. We summarize effects seen in both functional and structural parameters in the following.

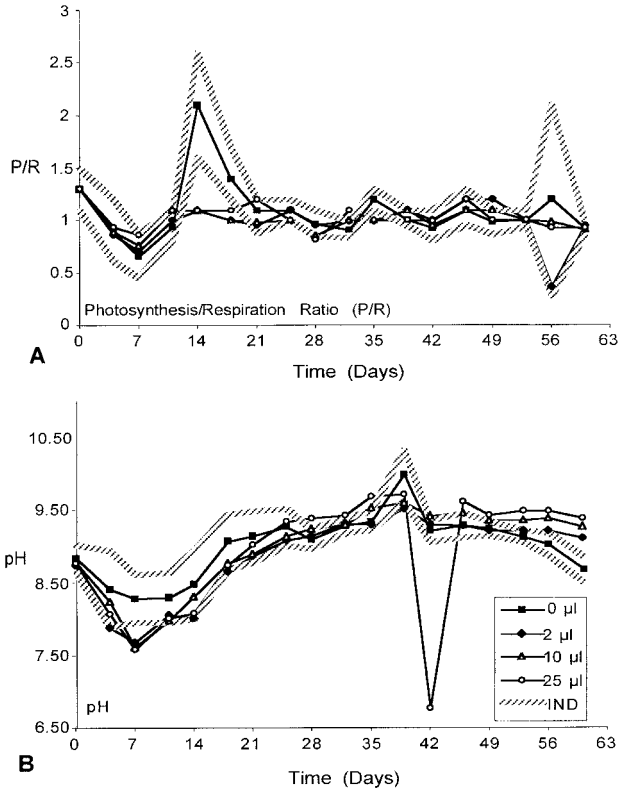

Fig. 2. Changes in the photosynthesis-to-respiration (P:LR) ratio (**A**) and pH (**B**) over time.

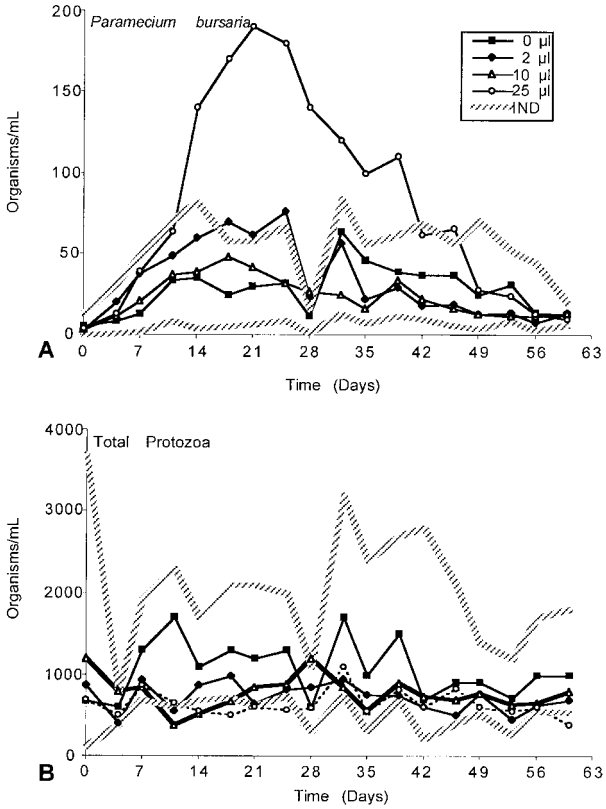

Fig. 3. Changes in *Paramecium bursaria* and total protozoa over time. The 25-µl treatment responded with a bloom of *P. bursaria* (**A**) that persisted during most of the experiment. Total protozoa (**B**) demonstrated an initial depression of protozoa in the two highest treatment groups early in the experiment. By day 28, no difference was apparent.

Functional parameters

The P:R ratio demonstrated treatment-related effects during experiment days 11 and 21 of the experiment (Fig. 2). As the P:R ratio increased in the nondosed treatments, it remained depressed in the replicates dosed with Jet-A. This effect is due to statistically significant reductions in the dissolved oxygen concentrations in the morning and before the lights were turned on. Both morning dissolved oxygen measurements, AMDO1 and AMDO3, exhibit this treatment-related effect between experiment days 11 and 21. All three dosed treatments exceeded the IND.

The functional parameter pH also showed treatment-related effects (Fig. 2). An initial depression compared to the nondosed treatment was observed on day 7. From experiment days 28 to 35, the highest treatment group consistently exceeded the IND. The excursion on day 42 is due to a low pH measurement for a replicate. At the end of the experiment, from experiment days 49 to 63, the dosed treatments moved outside the IND, in a concentration-related manner, with the highest treatment group furthest outside the IND. The lowest dosed group (2 µl) exceeded the IND on the last two measurement days.

Structural parameters

Most structural parameters fell inside the nondosed treatment IND at day 0 of the experiment. Thereafter, a variety of treatment effects were detectable during the experiment.

Total algae demonstrated no consistent treatment effects during the experiment, and neither did the individual algal species. All treatments had similar dynamics with a steady increase in total algal density until the end of the experiment. Bluegreen algae, principally *Anabaena* sp., were the dominant types at the end of the experiment. Only *Scenedesmus* sp. demonstrated a decrease in density outside the IND for the 25-µl treatment, from days 11 to 18. A large IND range and the corresponding variability of the *Scenedesmus* sp. parameter were apparent.

The densities of the protozoan *Paramecium bursaria* demonstrated treatment-related effects, especially in the 25-µl group (Fig. 3A). Throughout much of the test, *P. bursaria* numbers were four times as dense in the 25-µl treatment group as in the nondosed treatment. The 2-µl group also exceeded the IND from day 17 until day 24. By day 49, the numbers in all the treatment groups were within the IND. Total protozoa numbers did not reflect the response of *P. bursaria* (Fig. 3B). The 10- and 25-µl treatment groups reflected a statistically significant decrease in total protozoa from days 11 to 18 and from days 11 to 25, respectively. During the remainder of the experiment, total protozoa were within the IND.

Hyalella azteca was very rare or extinct in the 25-µl treatment replicates from day 25 until the end of the experiment. This concentration is considerably lower than the estimated LC20 value. No concentration-related effect was noted for the other treatments.

Total invertebrates showed a treatment-related decrease that persisted until the end of the experiment (Fig. 4). A decrease below the IND was apparent in the 25-µl treatment from day 25 until day 63, and the 10-µl treatment exhibited a decrease below the IND from day 49 until the end of the experiment.

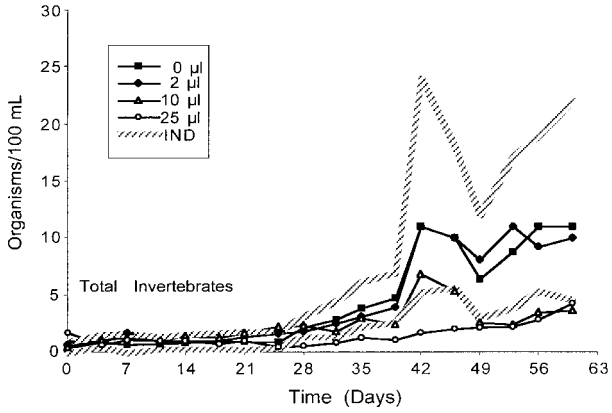

Fig. 4. Total invertebrate dynamics within the mixed flask culture (MFC). These numbers do not include protozoans. The highest treatment group is outside the intervals of nonsignificant difference (IND) after day 25, and the two highest treatment groups are outside the IND after day 49. *Hyalella azteca* was a dominant constituent of the community until after the bloom represented by day 42.

The 0- and 2-µl groups tracked very closely together from day 35 onward.

In the early stages of the experiment, *H. azteca* was the dominant invertebrate, reaching a peak on day 42. After this date, the number of amphipods decreased as other invertebrates began a rise. The distinction between the highest and lowest treatment groups at the end of the experiment was due to other members of the invertebrate assemblage increasing in number at different rates.

Multivariate analysis

Figure 5 shows the significance levels for the three multivariate tests used. All three methods demonstrate significant treatment clusters for both functional (Fig. 5A) and structural (Fig. 5B) parameters. Changes in significance are apparent from the plots based on comparisons of all three multivariate tests. Treatment-related patterns were not as apparent for the structural parameters as the functional parameters. This pattern

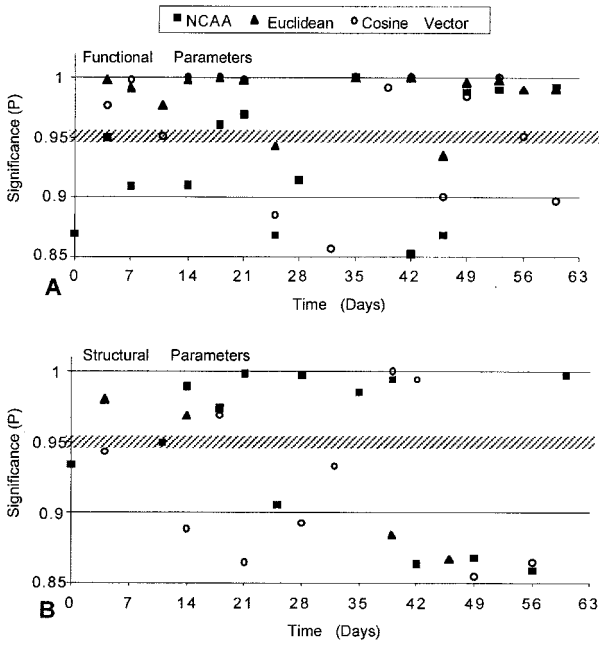

Fig. 5. Plots of significance over time of the Euclidean distance, cosine of the vector, and nonmetric clustering and association analysis (NCAA) multivariate analyses for both the functional (**A**) and structural parameters (**B**). Points above the 0.95 probability level are considered clusters significantly associated with a treatment group.

is especially apparent at the end of the microcosm experiment. All three methods were adept at detecting treatment-related patterns using the functional parameters. The NCAA was more effective in detecting effects using structural parameters, although only one significant result was detected after day 42 of the experiment.

Important variables identified by NCAA in determining treatment effect by day are shown in Table 2. The overall importance of the pH and the parameters reflecting oxygen demand (AMDO) corresponds well with univariate results portrayed previously. In particular, pH appears more important at

Table 2. Important variables ranked according to contribution for each sampling day as determined by nonmetric clustering for functional and structural parameters. Important variables are defined as those allowing a proportional reduction of error value in the clustering greater than or equal to 0.5. Hyphen between values denotes equal rank in the listing of importance in the clustering

Experiment day	Functional parameters	Structural parameters (organisms)
0	AMDO1, pH, Abs, AMDO3	Rotifers, *Chlorella* sp.-other diatoms, other green unicellular, ciliates
4	AMDO3, pH, AMDO1, PMDO2, Abs	*Chlorella* sp., ciliates, flagellates, *P. bursaria*
7	PH, AMDO1, AMDO3, PMDO2	Other diatoms, ciliates, *P. bursaria*
11	PH-PMDO2, AMDO1-Abs, AMDO3	Other diatoms, other bluegreens, flagellates, *Chlorella* sp.
14	PH, AMDO1, PMDO2, AMDO3	*Lyngbya* sp.-*P. bursaria-H. azteca, Scenedesmus* sp., copepods
18	PMDO2, PH, AMDO3, AMDO1	Flagellates, *P. bursaria, H. azteca, Scenedemus* sp., other green unicellular, rotifers
21	AMDO1, PH, AMDO3, PMDO2	Rotifers, *Scenedesmus* sp., other diatoms, *P. bursaria*
25	PH, AMDO1, PMDO2, Abs, AMDO3	*P. bursaria,* flagellates, *Scenedesmus* sp., other diatoms, *H. azteca,* copepods
28	PH, AMDO3, AMDO1	*P. bursaria, H. azteca, Scenedesmus* sp., *Lyngbya* sp., *Chlorella* sp., rotifers
32	AMDO1, AMDO3, PMDO2	Ciliates, rotifers-*Lyngbya* sp., other green unicellular
35	PH, PMDO2, Abs	*Ankistrodesmus* sp., *Scenedesmus* sp.-other diatoms-ciliates-other bluegreens-ostracod II
39	AMDO3, PMDO2, AMDO1, Abs	*P. bursaria,* other green unicellular, *H. azteca,* other diatoms
42	PMDO2, AMDO3, AMDO1, PH	Other green unicellular, ciliates, *H. azteca, Scenedesmus* sp.
46	AMDO3, AMDO1, PH	*P. bursaria,* copepods, ciliates, *Selenastrum* sp.
49	Abs, PH, AMDO1	*Scenedesmus* sp., flagellates, rotifers, *Nitzchia* sp.
53	PH-AMDO3, AMDO1	*P. bursaria, Selenastrum* sp., other green unicellular, other bluegreens
56	AMDO3, Abs, PMDO2-PH	*Scenedesmus* sp., *H. azteca,* other green unicellular
60	PH, AMDO1, AMDO3	Ciliates, rotifers, *Selenastrum* sp.

Table 3. Important variables according to success in determining treatment effect as determined by nonmetric clustering and association analysis (NCAA)

Variable	No. dates important in clustering
Functional variables	
PH	16
AMDO1	15
AMDO3	15
Top five most frequent structural variables	
Scenedesmus sp.	10
P. bursaria	9
Other green unicellular	8
Ciliates, rotifers, *H. azteca*	7 (tie)
Other diatoms	6

the experiment's beginning and end. A similar pattern is apparent in the univariate analysis (Fig. 2B). Table 3 ranks the important variables listed most often in the NCAA analysis. The order of the functional variables corresponds with the univariate data. In the clustering using structural variables, *Scenedesmus* sp. is ranked highest, although univariate data show a diversion outside the IND on only three sampling dates. The importance of *P. bursaria* in the clustering is reflected in the univariate results.

DISCUSSION

Evaluation of dosing technique

A simulation of natural freshwater sediment contamination from underlying sources was achieved. This was due to the large pore capacity of the overlying silica sand MFC sediment allowing access of the overlying water and detritus to the spiked layer. In addition, the incorporation of the powdered cellulose and chitin in the spiked sediment layer provided sorptive substrate for the hydrocarbons. This technique may also be valid for use with natural contaminated freshwater sediment either through the use of sediment dilution to obtain concentration–effect information [40,41] or simply with whole sediments from various contaminated sites.

The injection of neat Jet-A into the sediment allowed a reservoir of material that could be remixed or suspended in the water column as the microcosm was mixed for sampling. This allowed the toxicant to persist longer than in experiments using the water-soluble fraction of Jet-A and other turbine fuels [22,23,42].

Hyalella azteca *acute test and survival in the MFC*

Hyalella azteca went extinct during the experiment in the 25-μl treatment group. Other treatments demonstrated the viability of *H. azteca* in the MFC test system. This concentration of the highest treatment group is 20% that of the estimated acute EC20 from the 10-d toxicity test. The lack of survival of the population could be due to chronic effects, but survivorship in the 10-μl treatment was noted. Given the concentration of the toxicant, the extinction effect was likely due to a combination of chronic and indirect effects propagated in the microcosms of the highest treatment group. In this instance, an EC20 was not protective of the amphipod populations in the microcosms.

Comparison of the MFC and SAM experimental methods

The specific results of this MFC multispecies experiment are distinctive from those obtained by the SAM protocol using the same and related jet fuels [22,23,42]. The SAM experiments use a defined set of organisms inoculated into a nutrient-rich media. In our group's SAM experiments, an initial algal bloom occurred followed by a bloom in the daphnid populations. During the later stages of the experiment, ostracod populations increased, indicating a switch to a detritovore economy in the microcosms [43].

In contrast, the MFC begins with an established ecological community where the replicates are cross-inoculated until dosing. After dosing, two sets of effects occur: those caused by lack of further immigration and those due to the toxicant. Many species are present in the MFC, but the composition of the replicates is determined by the initial inoculum taken from the environment and the stochastic and deterministic events that occur during the establishment of the experimental replicates. In this experiment, a steady increase in the densities of algae in the replicates occurred, and an increase in the invertebrate population followed.

Both the protocols used the same media and sediment. The enumeration methods of the SAM were adapted for the smaller MFC test vessel. The univariate and multivariate data analysis techniques were identical to Landis et al.'s [22,23] experiments. In the following discussion, we compare the responses of the functional and structural variables, then evaluate each for persistence of effects.

Effects in comparison to the SAM experiments

The dynamics in the community structure between the two types of experiments were fundamentally different. However, the persistence of effects in each type of experiment was similar despite the fundamentally different approaches.

The P:R ratio and pH were observed in both experiment types. In both the SAM and the MFC, pH indicates toxicant impact. In the Jet-A SAM, the pH is above that of the nondosed reference. In the MFC, pH was initially lower, then at the end of the experiment the pH was significantly higher in a concentration–response pattern. In both experiments, the P:R ratio was significantly different in the beginning of the experiment but not at the end.

Structural responses to the toxicant were very different in the two experiments. No algae bloom occurred in the MFC as did with the SAM. One characteristic of turbine fuel impact in the SAM is the increase in algal populations caused by cropping of the daphnia herbivores by the toxicant. No such pattern occurs in the MFC. Changes do occur in the invertebrate dynamics of both experiment types, although the responses are different. In the SAM experiments, as the jet fuel degrades, the daphnia experience a bloom as the algae are ingested. Size and timing of the daphnid bloom are concentration dependent. Ostracod populations increase in the latter half of the SAM experiments, with the higher concentrations of jet fuel producing higher densities of ostracods. In the MFC experiments, invertebrates did show concentration-related impacts in the latter half of the experiment, but the higher concentrations resulted in lower invertebrate densities. The protozoa *Tetrahymena* in the SAM experiments were not sensitive measures of toxicant impacts. However, *P. bursaria* is in the MFC experiment.

In both the SAM and the MFC experiments, the clustering techniques detected treatment effects for the duration of the

experiment. In the Landis et al. [22,23] SAM experiments, the structural variables were used exclusively, although Landis [42] used both structural and functional variables. In MFC, we separated the functional and structural variables. Both sets of variables detected treatment-related effects. The structural variables were not as effective as the functional variables in detecting treatment groups.

Persistence of effects

Persistence of detectable effects is a characteristic of all the SAM and MFC experiments. Analytical chemistry of the water column in the SAM experiments demonstrated the rapid degradation of the water-soluble fraction. The MFC was dosed with the neat material, and some bound to the sediment to be released on mixing. Perhaps some of the observed effects in the MFC could be due to materials not detectable in the water column but bound to the components of the artificial sediment or the accumulated detritus.

The variables that allow the detection of treatment effects change over time in both types of experiments. Table 2 clearly shows these changes as demonstrated by NCAA for the MFC. Examination of the density plots also bears this fact. *Paramecium bursaria* was particularly sensitive to treatment group during the middle of the experiment. Invertebrates and pH were good indicators of treatment effects at the end. The change in the indicators of effects is noted in the SAM experiments as well and likely is a universal property of these and other ecological systems.

The persistence of these treatment effects, the changes in the indicators of these effects, and the repeatability of these properties in SAM experiments led to formulation of the community conditioning hypothesis [29,43]. This hypothesis states that ecological communities tend to preserve information about every event in their history, long after the disturbance occurred. The information can be carried in the structure of the community, in the population dynamics and structure of the constituent populations, and in the structure of their genomes. Corollaries to this are that ecological systems are complex, nonequilibrium systems and that the best measures of effects change over time [42].

As discussed previously, the MFC experiments demonstrated these properties. In the SAM experiments, the water-soluble fraction of the jet fuel degraded during the course of the experiment. In the MFC experiment, the toxicant could apparently be stored in the sediment for subsequent degradation and release into the water column. It is possible that the stressor remained in the system, at low and transient concentrations in the water column, and this residual could be directly causing the effects at the end of the experiment.

In this and other microcosm experiments, we routinely fail to find recovery as defined as concurrence with the nondosed treatment. Admittedly, the lowest concentration in this and other experiments tracks the nondosed treatments in most of the parameters. However, examination of the graphs for parameters such as pH and total invertebrates demonstrates a treatment-related effect late in the experiment. Total invertebrates exhibit two groups at the end of the MFC, the first group being comprised of nondosed and 2-μl treatments and the second group comprised of 10- and 25-μl treatments. Perhaps the variety of parameters that are examined, the statistical power of six replicates per treatment, and the utilization of multivariate techniques allow sufficient power to detect subtle changes in the treatment groups.

Does a failure to detect statistically significant differences in the parameters indicate that the ecological system has returned to its prestressed state? The answer has been demonstrated to be conclusively no by Kersting and Van Wijngaarden [12] using a very different multispecies system. The ditch microcosms were dosed twice with the herbicide Linuron (Crescent Chemical, Hauppauge, NY, USA). The functional parameters of the dosed systems were not different statistically from the nondosed treatment. However, on redosing, the previously dosed systems were much more resistant to the herbicide. The change in the sensitivity of the system indicates an important functional change in the system, although the measured variables did not reveal the alteration.

There are important implications of this and related experimental multispecies tests for field research. These microcosm experiments, because of their relatively low variability and number of true replicates, can find treatment related impacts at very low concentrations of materials. The statistical power found in the microcosm experiments should be much higher than that of a typical field study that, by its very nature, lacks replicability and has no control over many important variables. The lack of statistical power of field studies makes an illusion of recovery or no significant difference using conventional statistical techniques a likely outcome. Approaches used to date in most field research and in experimental microcosms, including ours, do not incorporate changes in system dynamics as an endpoint. Dealing with dynamic systems is difficult statistically with typical ecological data sets, although methods are under development [42].

Acknowledgement—We would like to thank the Institute of Environmental Toxicology and Chemistry support staff, notably D. Hennessy, A. Markiewicz, Mike Roze, L. Holmquist, and M.K. Moores. This research was supported by U.S. Air Force Office of Scientific Research Grant AFOSR-91-0291 DEF.

REFERENCES

1. Lyman WJ. 1984. Establishing sediment criteria for chemicals—Industrial perspectives. In Dickson KL, Maki AW, Brung WA, eds, *Fate and Effects of Sediment-Bound Chemicals in Aquatic Systems*. Pergamon, New York, NY, USA, pp 378–387.
2. Southerland E, Kravitz M, Wall T. 1992. Management framework for contaminated sediments (U.S. EPA Sediment Management Strategy). In Burton GA Jr, ed, *Sediment Toxicity Assessment.* CRC, Boca Raton, FL, USA, pp 341–370.
3. U.S. Environmental Protection Agency. 1987. An overview of sediment quality in the United States. EPA 905/9-88-002. Office of Water Regulations and Standards, Washington, DC.
4. Burton GA Jr. 1991. Assessing the toxicity of freshwater sediments. *Environ Toxicol Chem* 10:1585–1627.
5. Chapman PM. 1989. Current approaches to developing sediment quality criteria. *Environ Toxicol Chem* 8:589–599.
6. Cairns J Jr. 1986. The myth of the most sensitive species. *BioScience* 36:670–672.
7. Kimball KD, Levin SA. 1985. Limitations of bioassays: The need for ecosystem-level testing. *BioScience* 35:165–171.
8. Giddings JM. 1981. Laboratory tests for chemical effects on aquatic population interactions and ecosystem properties. In Hammons AS, ed, *Methods for Ecological Toxicology: A Critical Review of Laboratory Multispecies Tests.* Ann Arbor Science, Ann Arbor, MI, USA, pp 23–91.
9. Leffler JW. 1984. The use of self-selected, generic aquatic microcosms for pollution effects assessment. In Harris HH, ed, *Concepts in Marine Pollution Measurements.* Maryland SeaGrant College, University of Maryland, College Park, MD, USA, pp 139–158.
10. Taub FB, Read PL. 1982. Standardized aquatic microcosm protocol, Vol 2. 223-80-2352. Final Report. Food and Drug Administration, Washington, DC, USA.
11. Touart LW. 1988. Hazard Evaluation Division Technical Guidance

Document—Aquatic mesocosm test to support pesticide registrations. EPA 540/9-88/035. U.S. Environmental Protection Agency, Washington, DC, pp 1–41.
12. Kersting K, Van Wijngaarden PA. 1999. Effects of a pulsed treatment with the herbicide afalon (active ingredient Linuron) on macrophyte-dominated mesocosms. I. Responses of ecosystem metabolism. *Environ Toxicol Chem* 18:2859–2865.
13. Kedwards KJ, Maund SJ, Chapman PF. 1999. Community level analysis of ecotoxicological field studies: I. Biological monitoring. 1999. *Environ Toxicol Chem* 18:149–157.
14. Kedwards KJ, Maund SJ, Chapman PF. 1999. Community level analysis of ecotoxicological field studies: II. Replicated-design studies. *Environ Toxicol Chem* 18:158–166.
15. Van Den Brink PJ, Ter Braak CJF. 1999. Principal response curves: Analysis of time-dependent multivariate responses of biological community to stress. *Environ Toxicol Chem* 18:138–148.
16. Sparks TH, Scott WA, Clarke RT. 1999. Traditional multivariate techniques: Potential for use in ecotoxicology. *Environ Toxicol Chem* 18:128–137.
17. Van Wijngaarden RPA, Van den Brink PJ, Crum SJH, Oude Voshaar JH, Brock TCM, Leeuwangh P. 1996. Effects of the insecticide Dursbant 4E (active ingredient chlorpyrifos) in outdoor experimental ditches: I. Comparison of short-term toxicity between laboratory and field. *Environ Toxicol Chem* 15:1133–1142.
18. Landis WG, Markiewicz AJ, Matthews GB, Roze MJ, Matthews RA. 1998. Case study #2: An exploration of uncertainty in the determination of environmental toxicity. In Warren-Hicks WJ, Moore DRJ, eds, *Uncertainty Analysis in Ecological Risk Assessment.* SETAC, Pensacola, FL, USA, pp 226–236.
19. Landis WG, Markiewicz AJ, Matthews RA, Matthews GB. 2000. A test of the community conditioning hypothesis: Persistence of effects in model ecological structures dosed with jet fuel JP-8. *Environ Toxicol Chem* 19:327–336.
20. Landis WG, Matthews GB, Matthews RA, Sergeant A. 1994. Application of multivariate techniques to endpoint determination, selection and evaluation in ecological risk assessment. *Environ Toxicol Chem* 12:1917–1927.
21. Conquest LL, Taub FB. 1989. Repeatability and reproducibility of the standard aquatic microcosm: Statistical properties. In Cowgill UM, Williams LR, eds, *Aquatic Toxicology and Hazard Assessment,* Vol 12. STP 1027. American Society for Testing and Materials, Philadelphia, PA, pp 159–177.
22. Landis WG, Matthews RA, Markiewicz AJ, Matthews GB. 1993. Multivariate analysis of the impacts of the turbine fuel JP-4 in a microcosm toxicity test with implications for the evaluation of ecosystem dynamics and risk assessment. *Ecotoxicology* 2:271–300.
23. Landis WG, Matthews RA, Markiewicz AJ, Shough NA, Matthews GB. 1993. Multivariate analyses of the impacts of the turbine fuel Jet A using a microcosm toxicity test. *J Environ Sci* 2: 113–130.
24. Matthews GB, Matthews RA, Hachmoller B. 1991. Mathematical analysis of temporal and spatial trends in the benthic macroinvertebrate communities of a small stream. *Can J Fish Aquat Sci* 48:2184–2190.
25. Matthews RA, Matthews GB, Ehinger WJ. 1991. Classification and ordination of limnological data: A comparison of analytical tools. *Ecol Model* 53:167–187.
26. Matthews GB, Matthews RA, Landis WG. 1995. Nonmetric conceptual clustering in ecology and ecotoxicology. *AI Appl* 9:41–48.
27. Kersting K, Van den Brink PJ. 1997. Effects of the insecticide Dursbant 4E (active ingredient chlorpyrifos) in outdoor experimental ditches: III. Responses of ecosystem metabolism. *Environ Toxicol Chem* 16:251–259.
28. Warren-Hicks WJ, Moore DRJ. 1998. Issues and methods in uncertainty analysis. In Warren-Hicks WJ, Moore, DRJ, eds, *Uncertainty Analysis in Ecological Risk Assessment.* SETAC, Pensacola, FL, USA, p 48.
29. Landis WG, Matthews RA, Matthews GB. 1996. The layered and historical nature of ecological systems and the risk assessment of pesticides. *Environ Toxicol Chem* 15:432–440.
30. Lawton JH. 1996. The Ecotron facility at Silwood Park: The value of ''big bottle'' experiments. *Ecology* 77:665–669.
31. Drake JA, Huxel GR, Hewitt CL. 1996. Microcosms as models for generating and testing community theory. *Ecology* 77:670–677.
32. Carpenter SR. 1996. Microcosm experiments have limited relevance for community and ecosystem ecology. *Ecology* 77:677.
33. Van Geest GJ, Zwaardemaker NG, Van Wijngaarden RPA, Cuppen JGM. 1999. Effects of a pulsed treatment with the herbicide afalon (active ingredient Linuron) on macrophyte-dominated mesocosms. II. Structural responses. *Environ Toxicol Chem* 18: 2866–2874.
34. American Society for Testing and Materials. 1994. Standard guide for conducting sediment toxicity tests with freshwater invertebrates, E 1383-94. In *Annual Book of ASTM Standards,* Vol 11.04. Philadelphia, PA, pp 1196–1225.
35. American Society for Testing and Materials. 1991. Standard practice for the standardized aquatic microcosm: Fresh water. E 1366-91. In *Annual Book of ASTM Standards,* Vol 11.04. Philadelphia, PA, pp 1017–1051.
36. Good IJ. 1982. An index of separateness of clusters and a permutation test for its significance. *J Stat Comp Simul* 15:81–84.
37. Smith EP, Pontasch KW, Cairns J Jr. 1990. Community similarity and the analysis of multispecies environmental data: A unified statistical approach. *Water Res* 24:507–514.
38. Noreen EW. 1989. *Computer Intensive Methods for Testing Hypotheses.* Wiley-Interscience, New York, NY, USA.
39. Matthews GB, Hearne JW. 1991. Clustering without a metric. *IEEE Trans Pattern Anal Mach Intell* 13:175–184.
40. Giesy JP, Rosiu CJ, Graney RL, Henry MG. 1990. Benthic invertebrate bioassays with toxic sediment and pore water. *Environ Toxicol Chem* 9:233–248.
41. Landrum PF, Tsymbal VN, Nelson MK, Ingersoll DC, Gossiax DC, Burton GA, Brickson G. 1990. Sediment associated contaminant toxicity assessment by dilution experiments. *Abstracts,* 11th annual meeting, Society of Environmental Toxicology and Chemistry, Arlington, VA, USA, November 11–15, p 107.
42. Landis WG, JF McLaughlin. 2000. Design criteria and derivation of indicators for ecological position, direction and risk. *Environ Toxicol Chem* 19:1059–1065.
43. Matthews RA, Landis WG, Matthews GB. 1996. Community conditioning: An ecological approach to environmental toxicology. *Environ Toxicol Chem* 15:597–603.