United States
Environmental Protection
Agency

EPA

Research and Development

Sheboygan Confined Treatment Facility (CTF)
Bioremediation Workshop

Prepared for

Great Lakes National Program Office
Region V
U.S. Environmental Protection Agency

Prepared by

Environmental Research
Laboratory
Athens GA 30613

February 1992

Sheboygan Confined Treatment Facility (CTF) Bioremediation Workshop

On May 30, 1991, a workshop was convened by the EPA Environmental Research Laboratory, Athens, GA to review a bioremediation pilot study to be conducted at the Sheboygan Confined Treatment Facility (CTF). The objectives of the workshop were to :

1. Conduct a scientific review of the work plan for the pilot bioremediation study and provide recommendations for enhancing the experimental design.

2. Finalize a sampling plan for monitoring the CTF that would produce a statistically supportable documentation of the effectiveness of bioremediation.

3. Set up a plan for conducting a scientific evaluation of the monitoring data in order to determine loss of PCBs due to biodegradation as compared to losses resulting from other mechanisms.

Participants comprised representatives from Blasland and Bouck (site contractors), Michigan State University, Wisconsin Department of Natural Resources, EPA Region V, and the EPA Environmental Research Laboratory - Athens. A complete list of participants can be found in Appendix A.

Dr. Mark Brown of Blasland and Bouck Engineers started the meeting with a description of the bioremediation pilot study. The opening remarks were followed by an open discussion of the study. What follows is a summary of the workshop.

A CTF has been constructed on the premises of the Tecumseh Products Company, Sheboygan Falls, WI by Blasland and Bouck Engineering. The facility is constructed of steel sheet piling (105 ft x 135 ft x 10 ft) lined with welded sheets of high-density polyethylene into which are inlaid pipe grids for detection of leaks and injection of amendment. The structure is divided in half to permit testing of two treatments, and within each half, approximately one-fourth the total area has been segregated and reserved as a treatment control. In addition, several materials are being tested as adsorbents for PCBs in the outflow water.

Sediments were mechanically dredged from the Sheboygan River and transported to the facility in leak-proof containers. Although each load of sediment was not characterized before deposition in the facility, comparability of material in each treatment/control pair was attempted by delivering similar

proportions of sediment from each of the dredging sites. Cells 1 and 2 were considered to have similar sediment/water compositions; cells 3 and 4 also were considered to be similar. However, cells 3 and 4 may be different from cells 1 and 2. Cells 2 and 3 are approximately one third the size of cells 1 and 4, respectively. Cells 2 and 3 were meant to serve as controls for cells 1 and 4, respectively. The requirements of mechanical dredging preclude mixing of the sediments before deposition in the CTF. Thus, the heterogeneous nature of the material in the facility introduces high variance into measurements of baseline levels of contaminants and responses to treatments.

The sampling of sediments in the CTF cells has been limited to cell 4. Cell 4 was sampled by hand-held corers. A composite sampling approach was used to determine PCB concentrations. Composites were prepared from samples collected randomly from among a set of systematic samples taken from the cell. The composite samples were thoroughly mixed and then the mixture was subsampled for analysis. Material from seven cores were mixed to create seven composite samples. Only one subsample has been analyzed per composite sample thus far. The data so derived were utilized to estimate the mean PCB concentration in the cell, and the sample variance was calculated to measure variability. The mean PCB concentration of the composite samples was 325 +- 175 ppm. Cells 1, 2, and 3 were not sampled. This information was used to project the number of samples needed to detect 50% degradation in PCB levels.

Several areas within the river have been armored by Blasland and Bouck Engineering to keep PCB contamination in place and eliminate transport and dissolution as loss mechanisms for sediment PCBs. Four areas of sediment having PCB concentrations ranging from 25 to 270 ppm were armored, and removable sections were installed for access to the underlying sediments. The rate of degradation in the armored area may serve as an indication of <u>in situ</u> rates and as a point of comparison for rates in the CTF.

The CTF was constructed while bench-scale experiments were being conducted at the University of Michigan; the most promising treatments from the bench-scale work are to be implemented and tested in the pilot facility.

Aerobic experiments were conducted on partially dechlorinated sediments and anaerobic experiments were performed with contaminated sediments having low (10 ppm) and high (500 ppm) PCB concentrations. Results to date suggest that the major component of the original contaminant was Aroclor 1248 and that dechlorination is occurring at this contaminated area of concern. Data collected from 1989 to 1991 support this suggestion since a plot of the number of <u>ortho</u>- versus <u>meta</u>- and <u>para</u>-chlorines showed a shift away from the standard plot of Aroclor 1248. Dechlorination occurred at the <u>meta</u>- and <u>para</u>-positions of highly

chlorinated congeners which resulted in the accumulation of less chlorinated ortho-substituted congeners. A faster rate of dechlorination was observed in sediments with higher concentrations of PCBs than the less contaminated sediments. Only mono- and dichlorinated congeners were degraded in the aerobic experiments, but the higher chlorinated congeners were degraded in the anaerobic experiments.

The effect of amending anaerobic sediments (artificially contaminated with Aroclor 1242) with inorganic and organic nutrients also was investigated in bench scale experiments. Experimental treatments tested included mineral (RAMM medium) and secondary substrate (methanol and acetone) additions. There was little conclusive evidence that addition of organic substrate plus RAMM medium enhanced rates of PCB (500 or 10 ppm) dechlorination at an incubation temperature of 30 °C. The rate observed of dechlorination was not enhanced significantly by these two amendments, and an equivalent rate of dechlorination was observed in the control vessels without addition of organic substrates. Data from non-RAMM controls were not available. Non-RAMM control experiments were initiated in May 1991.

Future experiments are designed to increase the bioavailibility of sorbed PCBs. Surfactants, especially Triton X-705, will be used to increase the solubility of PCBs. Methods for transforming anaerobic sediments to aerobic conditions will be studied since the products of anaerobic dechlorination may serve as substrates for aerobic bacteria. In particular, the rate and amount of hydrogen peroxide needed to transform anaerobic to aerobic sediments will be investigated.

In discussions that followed the introduction by Dr. Mark Brown, three major areas were addressed: the sampling design for evaluating treatment options, the treatment options to be tested this year, and the alternative treatments that could be pursued.

Concern was voiced by an EPA statistician and others that composite samples obscured intersample variation and that the intracomposite variance had to be established (by subsampling) for valid comparisons of values.

A composite sampling approach probably can be used advantageously for estimating mean PCB concentrations in the CTFs, but its use must be treated carefully in order that underestimation of variances or standard errors is avoided. Underestimation could lead to declaring significant differences where they do not really exist. It is important to realize that there are different sources of variability in this experiment and there has been no attempt to estimate or compare them. This means that such composites cannot be used to fine-tune a proposed design, as might be required when determining how many individual samples should be combined per composite sample, how many

composite samples should be formed, how many subsamples should be collected from each, whether the mixing process is adequate, and whether the laboratory-based 10-gram subsampling is acceptable. A study to address all of these questions apparently would be prohibitively expensive. (A note on composite sampling can be found in Appendix B.)

Because details of the computations were not available at the meeting, it is not possible to determine whether such calculations have been correctly implemented. This should be checked thoroughly inasmuch as small sample sizes have been suggested as being adequate. The number of proposed composite samples (five) does seem to be woefully small. The idea of using a 50% reduction in PCB levels as the tolerable difference (i.e., a difference to be detected with a given confidence level) seems to be too unrefined. It is a scientific question as to how long a 50% reduction is likely to take, and whether it is of interest to detect smaller changes. If smaller changes are of interest, there will be a direct impact on the number of samples required.

If it is possible, available data should be utilized to estimate variance components and/or standard errors, and simulations should be conducted to investigate the proposed sampling strategies. This modification was subsequently incorporated into the sampling design. Subsequent to the meeting Dr. Mark Brown proposed (Appendix C, Dr. Rudy Parrish's response Appendix D) a sampling plan that would use 5 seven-core composites from each of the cells. For one composite in each cell, triplicate subsamples would be collected. The relative standard deviations would be compared to determine whether an alternative composite sampling design is needed.

Appendix B contains applicable expressions for the variance of the estimate of the cell mean (Eq.4). This expression involves four variance components. Fortunately, it is not necessary to estimate these components individually in order to obtain an estimate of the variance of the estimator; however, the computation is facilitated by use of analysis of variance calculations whenever the data are balanced. Balanced, here, means that the same number of subsamples has been taken from each composite. In the unbalanced case, it is not clear exactly what the expressions should be, although they can probably be derived. If the cores are sufficiently large relative to the subsamples, then "S" will be large relative to "s", so that the quantity "(S-s)/(S-1)" is approximately the same as "S/(S-1)". If this is, in fact, the case, then (EMS_B - EMS_W) can be used to produce an estimate of the second term of Eq.4 that is associated with the composite-related variance components. It seems that the proposal is to calculate the variances of the cell means for subsampling a single time versus subsampling three times, and then compare the results. While that might be appropriate from a general point of view, it might be better to attempt to estimate

the variance components for: (1) between increments, (2) measurement error, and (3) the nonlinear combination of components due to imperfect compositing. The ANOVA-based approach might work well here if the design can be balanced.

The unfortunate aspect of this approach is that design questions cannot be rigorously addressed without a knowledge of all of the individual variance components. Estimation of these requires a more complex (and expensive) design. Using the procedure above (with the simplifying assumption), one could investigate effects of differing numbers of subsamples, since "s" is a divisor in the second term of Eq.4. Similarly, one could investigate the effects of differing numbers of subsamples on the number of composite samples "r". The effect of changing the number of samples per composite, however, cannot be examined in the same manner because "n" occurs nonlinearly.

It would also be helpful to consider unbalanced method in terms of the variance components. Without that, it seems that a better approach would be to utilize an equal number of subsamples from each composite. Of course, it might also be possible to derive appropriate expressions for the unbalanced case.

There are several features of the design of the CTFs that could have an impact on any experimental study that is conducted at the site. The CTFs are constructed asymmetrically to provide four cells, each of a different size.

Statistical concerns include the following.

(1) There might be a CTF-size effect that could interfere with treatment-control or treatment-treatment comparisons.

(2) The composition of cells 1 and 2 and that of 3 and 4 might not be sufficiently similar to justify treating the smaller cells as controls for the corresponding larger cells.

(3) The sediments in the cells were not mixed prior to being added, so that high spatial variability could be expected. There are also varying sediment depths. This impacts the level of sampling required for any given comparison or parameter estimation.

(4) Having only four cells limits capability to use experimental designs that involve replication. A suggestion here is that perhaps cells 1 and 2 and cells 3 and 4 may be considered to form two blocks with two homogeneous experimental units (i.e., cells) in each.

One further aspect of the design that should be mentioned is that the effectiveness of a given amendment option will be measured primarily on the basis of a change in concentration in a

single cell over some time period, perhaps compared to similar observations in a corresponding control cell. There is only one experimental unit per option, so that any observed differences may, in fact, be confounded with the composition of the cell being used. There is no replication, which is needed to estimate the experimental error for testing differences among treatments and, thus, to make sound inferences.

It might be possible to improve on the limitation of having only four compartments (cells) in the CTF. Cells 1 and 4 are large enough that they could be subdivided to form a total of eight cells in the facility. If the sediment composition in these could be distributed more uniformly among such cells, there would be wider latitude available for utilizing different experimental designs involving more replication (thus, higher power) and more treatments.

Several concerns were raised about the study's capability to evaluate the potential for biodegradation of endogenous PCBs and also to develop new techniques that will enhance bioremediation of the Sheboygan River and Harbor area of concern. These studies are currently going on in the laboratory of Dr. Timothy Vogel at the University of Michigan.

The laboratory experiments with Sheboygan sediments were conducted at 30°C. The justification for choosing this temperature was probably to obtain somewhat faster dechlorination rates to allow more rapid evaluation of the effects of different amendments, etc. The optimum temperature for dechlorination by the upper Hudson River microorganisms, however, is between 20 and 25°C, and <u>para</u>-dechlorinating activity is lost at 30°C. Thus, conducting laboratory experiments with the Sheboygan sediments at 30°C does not necessarily yield any advantage (in terms of more rapid assays), and has the obvious disadvantage that the results are not directly transferable to the field. While the Blasland and Bouck representatives did not have information on actual environmental temperatures, the sediments in the CTF are undoubtedly lower than 30°C. Average summer and winter temperatures for different depths in the CTF should be determined and future laboratory experiments conducted at comparable temperatures ranges.

Blasland and Bouck Engneering has so far determined oil and grease levels for only a few of the Sheboygan sediment samples. Oil and grease have been shown to inhibit the aerobic degradation of PCBs by reducing their bioavailability. The same is likely true for anaerobic dechlorination. Results from Dr. John Quensen's laboratory at Michigan State University suggest that oil and grease levels above approximately 1,000 ug/g sediment (dry weight) begin to decrease the rate of dechlorination. Levels of 10,000 ug/g strongly inhibit dechlorination.

Variations in oil and grease concentrations may account for the various extents of *in situ* dechlorination observed in the Sheboygan sediments. The higher levels could limit dechlorination within the CTF. Sediments should be sampled routinely for oil and grease. If levels are, in fact, high enough to limit the dechlorination process, greater overall rates of PCB transformation may be achieved by cycling between aerobic and anaerobic conditions. The oil and grease should be degraded during the aerobic cycles.

Several treatment options to address the question of availability of sediment-sorbed PCBs were discussed. Sediment organic matter, especially sediment oil and grease, strongly sorb PCBs, thus decreasing the rate of degradation. It was suggested that these organic materials might be decreased during an aerobic degradation period preceding the anaerobic treatment (e.g., 2 months aerobic followed by 8 months anaerobic). The sensitivity of the dechlorinating microorganisms to such cycling would need to be evaluated, but at least some of them are now known to be spore formers. Moreover, brief (30 minute) aeration of Hudson River inoculum did not diminish its subsequent activity under anaerobic conditions.

The mechanics of altering the aeration status of the CTF sediments were discussed; the plumbing arrangement for the facility makes drainage difficult. Moreover, drainage water would have to be treated before discharge. The possibility of adding dissolved oxidants, including peroxides, as a means of shifting oxidation status was also addressed. The final selection of treatment regimes for the CTF will be made pending further discussions among representatives of Blasland and Bouck Engineering, EPA scientists from the Office of Research and Development, and EPA Region V representatives.

The aerobic degradation of PCBs is a co-metabolic process. The responsible microorganisms are induced with biphenyl and in some cases with mono-chlorobiphenyls. An attractive feature of an anaerobic-aerobic biotreatment sequence is that levels of mono-chlorobiphenyls may be achieved anaerobically that can induce the co-metabolism of other congeners when conditions become aerobic.

The Sheboygan sediments, however, do not appear to accumulate the high levels of mono-chlorobiphenyls required. Thus aerobic degradation of the dechlorination products, in the absence of another inducer such as biphenyl, is likely to be quite limited. A possible solution is to inoculate the sediments with microorganisms that produce more monochlorobiphenyls, such as upper Hudson River microorganisms. This, however, may not be easily achieved. Drs. Vogel and Nies from the University of Michigan attempted such an inoculation by mixing Hudson River and Sheboygan sediments. They found this to have no effect on the

extent of dechlorination. Similar results were obtained by Dr. John Quensen by mixing inocula from the upper Hudson River and Silver Lake. The Silver Lake microorganisms normally produce high levels of 2,4,2',4'-tetrachlorobiphenyl from the dechlorination of Aroclor 1260. Rather than dechlorinating the 2,4,2',4'-PCB produced by the Silver Lake microorganisms, the upper Hudson River microorganisms apparently inhibited the activity of the Silver Lake microorganisms.

Further attempts should be made at inoculating the Sheboygan sediments to achieve higher levels of mono-chlorobiphenyls. Sediments from sites other than the upper Hudson River should be tried. More work needs to be done to further the understanding of compatibility and incompatibility of different PCB dechlorinating activities.

Blasland and Bouck engineers proposed several criteria besides total PCB concentration for assessing the progress of biotreatment in the CTF and armored areas in the river. These included toxic equivalency, total concentration of key dechlorination products, and average number of meta- and para-chlorines per biphenyl. These criteria emphasize different parameters, but are actually different ways of reducing the raw data consisting of congener-specific PCB analysis. The first two criteria are related to health and environmental hazards, whereas the last two are related to the extent of anaerobic dechlorination. All the above criteria should be considered (and all raw data retained for further possible evaluation after the experiment is finished).

It appears that different criteria should be applied to the CTF and the armored areas. Armoring is essentially a technique for reducing risk (by decreasing the chance of exposure). Decreases in toxic and bioaccumulative congeners would further reduce risk, and should be the major criteria applied to the armored areas. In the case of the CTF, there is the additional goal of reducing total PCB concentrations. This means creating aerobic conditions at the most effective time. Therefore, criteria related to the progress of anaerobic dechlorination are also pertinent. The observed increase in the proportion of mono-chlorinated biphenyls is probably most important as these congeners may induce the aerobic co-metabolism of other PCBs still present as indicated previously. One could go a step further than Blasland and Bouck engineers proposed and, based on the relative degradabilities reported in the literature for the various congeners remaining, calculate the expected decrease in total concentration to be expected if a switch were made to aerobic conditions.

PCB degradation rates appear to be slightly higher in situ than in the CTF, although, as mentioned, the values from the armored areas do not reflect the overall variability of sediment

rates, which may not be statistically different from the CTF rates. Given the few openings and the small areas that are intensively sampled at the armored sites, the data from these areas must be interpreted cautiously.

In summary the CTF provides a valuable outside laboratory for investigating technologies for the biological treatment of PCBs. One should be cautioned, however, that the selection of an appropriate sampling design is critical for the development of sufficient data to adequately evaluate each treatment option. Because of the size and construction of the facility, side-by-side testing of technologies would be difficult. Current research would suggest that several treatment options are available for testing. These options include addition of inorganic nutrients, organic nutrients, aerobic PCB-degrading bacteria, and the use of alternating aerobic/anaerobic systems. The proper selection of test conditions should result from the careful analysis of laboratory data.

Appendix A. List of Participants

Rebecca Adams
Technology Applications, Inc.
c/o Environmental Research Laboratory
U.S. EPA
Athens GA 30613-0801
404-546-3517

Rochelle Araujo
Environmental Research Laboratory
U. S. EPA
Athens GA 30613-0801
404-546-3468

Mark Brown
Blasland & Bouck Engineers, P.C.
6723 Towpath Road, Box 66
Syracuse NY 13214
315-446-9120

Dave Cowgill
Great Lakes National Program Office
230 South Dearborn
Chicago, IL 60604
312-353-3576

Bonnie Eleder
Region V, US EPA
230 South Dearborn Street
Chicago, IL 60604
312-886-4885

Terry Evanson
WI Department of Natural
Resources
101 S. Webster
PO Box 7921
Madison, WI 53707
608-266-0941

Dawn Foster
Blasland & Bouck Engineers, P.C.
6723 Towpath Road, Box 66
Syracuse NY 13214
325-446-9120

Jack Jones
Environmental Research Laboratory
U.S. EPA
Athens GA 30613-0801
404-546-3468

Rudy Parrish
Computer Sciences Corp.
c/o Environmental Research Laboratory
U.S. EPA
Athens GA 30613-0801
404-546-3604

John Quenson
Department of Crop and Soil Sciences
Michigan State University
East Lansing MI 48823-1325
517-355-9285

John Rogers
Environmental Research Laboratory
U.S. EPA
Athens GA 30613-0801
404-546-3592

Robin Schmidt
WI Department of Natural
Resources
101 S. Webster
PO Box 7921
Madison, WI 53707
608-267-7569

Patricia Van Hoof
University of Georgia
c/o Environmental Research Laboratory
U.S.EPA
Athens GA 30613-0801
404-546-3349

Daniel Wubah
University of Georgia
Environmental Research Laboratory
U.S. EPA
Athens GA 30613-0801
404-546-3182

Appendix B. A Note on Composite Sampling

Suppose $r \times n$ equal-sized portions of sediment material (increments) are drawn randomly from a cell. From these, r composites are formed by randomly partitioning the set of increments into r subsets of n increments each and physically mixing the increments in each subset. Then, s subsamples are drawn randomly from each composite, and t analyses are run on each subsample. Each composite is assumed to contain enough material for S subsamples. (If the subsample size were the same as the increment size, then $S = n$.)

To mathematically represent the variables involved in the composite sampling model, the following notation is defined:

a_{ijl} = the proportion of the jth subsample from the ith composite that comes from the lth increment in that composite,

X_{ijl} = the concentration value associated with the portion of the lth increment that appears in the jth subsample from the ith composite.

Each variable a_{ijl} has mean μ_a and variance $\sigma_{a^2}^2$. The variables X_{ijl} have equal means μ_x and variances $\sigma_x^2 + \sigma_w^2$, where σ_x^2 is the between-increment variance and σ_w^2 is the within-increment variance subsequent to compositing and subsampling.

In different composites, the variables a_{ijl} are uncorrelated, but they are correlated within the same increment, in the same subsample, and in different increments and subsamples within a composite. Similarly, the variables X_{ijl} are correlated within the same increment, in the same subsample, and in different increments and subsamples within a composite. Elder *et al.* (1980) provided explicit expressions for these covariances.

The concentration value associated with the jth subsample from the ith composite is based on contributions from all increments in that composite. It can be expressed as the weighted mean

$$y_{ij} = \sum_{l=1}^{n} a_{ijl} X_{ijl} \, . \qquad (1)$$

The value observed in the laboratory from the kth analysis, however, is

$$z_{ijk} = y_{ij} + e_{ijk} \quad (2)$$

where e_{ijk} is the measurement error associated with the kth test on the jth subsample from the ith composite. The variance of e_{ijk} is defined as σ_e^2. If the analysis procedure requires that additional subsampling be done, then an additional variance component should be incorporated.

The estimator of μ_x for composite sampling is

$$\bar{z} = \sum_{i=1}^{r} \sum_{j=1}^{s} \sum_{k=1}^{t} z_{ijk}/rst \; . \quad (3)$$

As a result of the covariance structures for the a_{ijl} and X_{ijl} variables, the variance for this estimator involves four individual variance components. Elder et al. (1980) provided

$$Var(\bar{z}) = \sigma_x^2/(rn) + [(S-s)/(S-1)][\sigma_w^2/n + n\sigma_a^2(\sigma_x^2 + \sigma_w^2)]/(rs) + \sigma_e^2/(rst) \; . \quad (4)$$

Also, they established that the expected mean squares (between groups, within groups, and for error) associated with the analysis of variance for the composite sample groups are:

$$EMS_B = (st/n)\sigma_x^2 + [t(S-s)/(S-1)][\sigma_w^2/n + n\sigma_a^2(\sigma_x^2 + \sigma_w^2)] + \sigma_e^2 \quad (5)$$

$$EMS_W = [tS/(S-1)][\sigma_w^2/n + n\sigma_a^2(\sigma_x^2 + \sigma_w^2)] + \sigma_e^2 \quad (6)$$

$$EMS_E = \sigma_e^2 \quad (7)$$

Clearly, the variance of the estimator $\bar{z}$ is equivalent to the value $EMS_e/(rst)$. Thus, the standard error of the estimate of μ_x can be obtained from a standard analysis of variance computation.

References

Brown, G. H. and N. I. Fisher. 1972. Subsampling a mixture of sampled material. Technometrics 14(3): 663-668.

Elder, Robert S., William O, Thompson, and Raymond H. Myers. 1980. Properties of composite sampling procedures. Technometrics 22(2): 179-186.

Gilbert, Richard O. 1987. Statistical Methods for Environmental Pollution Monitoring, Van Nostrand Reinhold Company, New York. 320 pp.

Parrish, R. S., G. O. Ware, C. N. Smith, and P. A. Banks. 1990. Composite sampling techniques for determining pesticide concentrations. Proceedings of the KSU Conference on Applied Statistics in Agriculture, Manhattan, KS.

Rhode, Charles A. 1976. Composite sampling. Biometrics 32, 273-282.

APPENDIX C

BLASLAND & BOUCK ENGINEERS, P.C.
ENGINEERS & GEOSCIENTISTS
6723 Towpath Road, Box 66, Syracuse, New York 13214 (315) 446-9120
FAX: (315) 449-0017

June 13, 1991

Rudy Parrish
Environmental Research Laboratory
U.S. EPA
Athens, GA 30613

Re: Bioremediation Pilot Study Work Plan

File: 176.06 #2

Dear Mr. Parrish:

We appreciate your review and discussion of sampling design as contained in the Sheboygan River and Harbor Bioremediation Pilot Study Work Plan. At the recent workshop in Atlanta it appeared that you and I reached an understanding that composite sampling was an acceptable approach. It was clear, however, that your acceptance of the approach was contingent upon an explicit assessment of the errors associated with subsampling the composite samples. This assessment may be useful in revising our sampling approach if the errors are large relative to the errors associated with our current methods. If the errors are large, we would conduct a numerical analysis to reassess the sampling requirements balancing the number of seven-core composite samples against the number of composite subsamples to minimize costs for a specified hypothesis-testing power.

We are not sure that some of the other workshop attendees are as yet comfortable with use of composite samples for gauging the overall progress of the experiment(s). A comment which was repeatedly offered is that our compositing scheme obscured the real spatial variability of PCB concentrations within the cell and that this variability needs to be explicitly considered. We do not dispute the first contention. On the contrary, the intended purpose of compositing was to cope with the expected high degree of spatial variability and the expense of analyzing large numbers of samples.

For the benefit of others, we are providing as an attachment, an example of how one might approach a sampling design from error estimates for various contributors. The approach illustrated by Neptune et al. can be compared and contrasted to our approach in developing a sampling design. The fundamental objective of both approaches was to produce a sampling design with an acceptable power for hypothesis testing. The essential difference is that by estimating the individual errors, the EPA authors could optimize the number of individual samples per composite sample and the numbers of

composite samples per area (note that sampling errors associated with subsampling of composite samples were ignored for composites of five or fewer samples) to produce the most cost-effective design. In contrast, we presumed that the spatial variability would be high and made an educated guess as to the number of cores required per composite sample. Based upon the actual distribution of PCB results for the composite samples, we then back calculated the requisite number of composite samples for a specified power. Having obscured the actual spatial variability (in this case, the distribution of average PCB concentrations in individual cores), we have no basis for altering the number of cores within a composite sample. So, for the time being, we must stay with the seven-core composites.

For the upcoming sampling, we propose to collect five seven-core composite samples from each of the four cells. In addition, we propose to submit triplicate subsamples of one composite sample from each of the three cells. The relative standard deviations of each of these composites would reflect not only our field errors in subsampling composite samples, but also subsampling errors by the laboratory and subsequent analytical errors. If these relative standard deviations are comparable to the relative standard deviations among composite samples within a cell, we would reconsider our sampling design based upon a numerical analysis (computer simulation of error distributions using random normal deviates). We would potentially recommend a revised balance of composite samples and subsamples of composite samples to minimize the costs for an acceptable hypothesis-testing power.

Please call me if you wish to discuss this. We look forward to continued cooperation with EPA's ARCS program.

Very truly yours,

BLASLAND & BOUCK ENGINEERS, P.C.

Mark P. Brown, Ph.D.
Manager

MPB/nlg
Attachment

cc: Bonnie L. Eleder, Environmental Protection Agency
Robin R. Schmidt, Wisconsin Department of Natural Resources
Dawn S. Foster, P.E., Blasland & Bouck Engineers, P.C.
John Rogers, Ph.D., Environmental Protection Agency
Chad Jafvert, Ph.D., Environmental Protection Agency

APPENDIX D

Computer Sciences Corporation
c/o U. S. Environmental Protection Agency
College Station Road
Athens, GA 30613-7799
Phone: 404-546-3604

22 July 1991

Mark Brown, Ph.D.
Blasland & Bouck Engineers, P.C.
6723 Towpath Road, Box 66
Syracuse, NY 13214

Dear Mark:

Since talking with you a few weeks ago, I spoke briefly with John Rogers about the Sheboygan project and your letter to me of June 13. Earlier, I provided John with summary comments pertaining to the May 30th meeting in Atlanta; enclosed is a copy of that memorandum for your reference. It had been my understanding that John would be communicating those and other comments to you.

The following remarks address your proposed work, as described in your letter. I realize that you may already have begun this work.

Your plan indicates that you will use 5 seven-core composites from each of the cells, and for one composite in each cell, you will collect triplicate subsamples. Then, compare relative standard deviations to determine whether an alternative composite sampling design is needed. The enclosed "Note on Composite Sampling" contains applicable expressions for the variance of the estimate of the cell mean (Eq.4). This expression involves four variance components. Fortunately, it is not necessary to estimate these components individually in order to obtain an estimate of the variance of the estimator; however, the computation is facilitated by use of analysis of variance calculations whenever the data are balanced. Balanced, here, means that the same number of subsamples has been taken from each composite. In the unbalanced case, it is not clear exactly what the expressions should be, although they can probably be derived. If the cores are sufficiently large relative to the subsamples, then "S" will be large relative to "s", so that the quantity "(S-s)/(S-1)" is approximately the same as "S/(S-1)". If this is, in fact, the case, then (EMS_B - EMS_w) can be used to produce an estimate of the second term of Eq.4 that is associated with the composite-related variance components. It seems that your proposal is to calculate the variances of the cell means for subsampling a single time versus subsampling three times, and then compare. While that might be appropriate from a general point of view, it might be better to attempt to estimate the variance components for: (1) between increments, (2) measurement error, and (3) the nonlinear combination of components due to imperfect compositing. The ANOVA-based approach might work well here if the design can be made balanced.

The unfortunate aspect of this is that design questions cannot be rigorously addressed without a knowledge of all of the individual variance components. Estimation of these requires a more complex (and expensive) design. Using the procedure above (with the simplifying assumption), one could investigate effects of differing numbers of subsamples, since "s" is a divisor in the second term of Eq.4. Similarly, for the number of composite samples "r". The effect of changing the number of samples per composite, however, cannot be examined in the same manner because "n" occurs nonlinearly.

Your procedure, I think, would be comparing two expressions based on Eq.4. It would be helpful to consider your method in terms of the variance components. Without that, it seems to me that a better

approach would be to utilize an equal number of subsamples from each composite. Of course, it might also be possible to derive appropriate expressions for the unbalanced case.

Hopefully, these comments will be of interest and of use to you. Please let me know if you would like to discuss this further.

Cordially yours,

Rudolph S. Parrish, Ph.D.
Statistician

[redacted]

June 10, 1991

To: John Rogers

Fm: Rudy Parrish

Re: Sheboygan Bioremediation Meeting Held in Atlanta May 30, 1991

As you requested, this is to summarize my impressions of the Sheboygan bioremediation work discussed at the meeting in Atlanta on May 30. These remarks are limited to statistical aspects of the experimental design proposed for the pilot study.

Design of confined treatment facility (CTF)

There are several features of the design of the CTFs that could have an impact on any experimental study that is conducted at the site. The CTFs are constructed asymmetrically to provide four cells, each of a different size. It was indicated that cells 1 and 2 have generally the same sediment/water composition, and similarly for cells 3 and 4. Cell 2 is of a size that is in approximately a 1:3 ratio to cell 1. Similarly, the size of cell 3 is in a 1:3 ratio to that of cell 4. Cells 2 and 3 were meant to serve as controls for cells 1 and 4, respectively.

Statistical concerns include the following.

(1) There might be a CTF-size effect that could interfere with treatment-control or treatment-treatment comparisons.

(2) The composition of cells 1 and 2 and that of 3 and 4 might not be sufficiently similar to justify treating the smaller cells as controls for the corresponding larger cells.

(3) The sediments in the cells were not mixed prior to being added, so that high spatial variability could be expected. There are also varying sediment depths. This impacts the level of sampling required for any given comparison or parameter estimation.

(4) Having only 4 cells limits capability to use experimental designs that involve replication. A suggestion here is that perhaps cells 1 and 2 and cells 3 and 4 may be considered to form two blocks with two homogeneous experimental units (i.e., cells) in each.

Sampling designs

The sampling of sediments in the CTF cells appears to have been limited to cell 4. The sampling approach has involved composite sampling techniques wherein several samples are collected randomly from among a set of systematic samples taken from the cell, the samples are combined physically, and then the mixture is subsampled. Only one subsample has been utilized per composite sample thus far. The data so derived were utilized to estimate the mean concentration in the cell, and the sample variance was calculated to measure variability. This information was used to project sample sizes needed to detect 50% degradation in PCB levels.

Because details of the computations were not available at the meeting, it is not possible to determine whether such calculations have been correctly implemented. This should be checked thoroughly inasmuch as small sample sizes have been suggested as being adequate. The number of proposed composite samples (five) does seem to be woefully too small. The idea of using a 50% reduction in PCB levels as the tolerable difference (i.e., a difference to be detected with a given confidence level)

seems to be too unrefined. It is a scientific question as to how long a 50% reduction is likely to take, and whether it is of interest to detect smaller changes. If smaller changes are of interest, there will be a direct impact on the number of samples required.

A composite sampling approach probably can be used advantageously for estimating means in the CTFs, but its use must be treated carefully in order that underestimation of variances or standard errors is avoided, which could lead to declaring significant differences where they do not really exist. It is important to realize that there are different sources of variability in this experiment and there has been no attempt to estimate or compare them. This means that such components cannot be used to fine-tune a proposed design, as might be required when determining how many individual samples should be combined per composite sample, how many composite samples should be formed, how many subsamples should be collected from each, whether the mixing process is adequate, and whether the laboratory-based 10-gram subsampling is acceptable. A study to address all of these questions apparently would be prohibitively expensive. (A note on composite sampling is attached.)

If it is possible, available data should be utilized to estimate variance components and/or standard errors, and simulations should be conducted to investigate the proposed sampling strategies.

Modification of the CTF

One further aspect of the design that should be mentioned is that the effectiveness of a given amendment option will be measured primarily on the basis of a change in concentration in a single cell over some time period, perhaps compared to similar observations in a corresponding control cell. There essentially is only one experimental unit per option, so that any observed differences may, in fact, be confounded with the composition of the cell being used. There is no real replication, which is needed to estimate the experimental error for testing differences among treatments and, thus, to make sound inferences.

It might be possible to improve on the limitation of having only four compartments (cells) in the CTF. Cells 1 and 4 are large enough that they could be subdivided to form a total of eight cells in the facility. If the sediment composition in these could be distributed more uniformly among such cells, there would be wider latitude available for utilizing different experimental designs involving more replication (thus, higher power) and more treatments.

A Note on Composite Sampling

Suppose $r \times n$ equal-sized portions of sediment material (increments) are drawn randomly from a cell. From these, r composites are formed by randomly partitioning the set of increments into r subsets of n increments each and physically mixing the increments in each subset. Then, s subsamples are drawn randomly from each composite, and t analyses are run on each subsample. Each composite is assumed to contain enough material for S subsamples. (If the subsample size were the same as the increment size, then $S = n$.)

To mathematically represent the variables involved in the composite sampling model, the following notation is defined:

a_{ij} = the proportion of the jth subsample from the ith composite that comes from the lth increment in that composite,

X_{ij} = the concentration value associated with the portion of the lth increment that appears in the jth subsample from the ith composite.

Each variable a_{ij} has mean μ_a and variance σ_a^2. The variables X_{ij} have equal means μ_x and variances $\sigma_x^2 + \sigma_w^2$, where σ_x^2 is the between-increment variance and σ_w^2 is the within-increment variance subsequent to compositing and subsampling.

In different composites, the variables a_{ij} are uncorrelated, but they are correlated within the same increment, in the same subsample, and in different increments and subsamples within a composite. Similarly, the variables X_{ij} are correlated within the same increment, in the same subsample, and in different increments and subsamples within a composite. Elder *et al.* (1980) provided explicit expressions for these covariances.

The concentration value associated with the jth subsample from the ith composite is based on contributions from all increments in that composite. It can be expressed as the weighted mean

$$y_{ij} = \sum_{l=1}^{n} a_{ijl} X_{ijl} . \tag{1}$$

The value observed in the laboratory from the kth analysis, however, is

$$z_{ijk} = y_{ij} + e_{ijk} \tag{2}$$

where e_{ijk} is the measurement error associated with the kth test on the jth subsample from the ith composite. The variance of e_{ijk} is defined as σ_e^2. If the analysis procedure requires that additional subsampling be done, then an additional variance component should be incorporated.

The estimator of μ_x for composite sampling is

$$\bar{z} = \sum_{i=1}^{r} \sum_{j=1}^{s} \sum_{k=1}^{t} z_{ijk} / rst . \tag{3}$$

As a result of the covariance structures for the a_{ijl} and X_{ijl} variables, the variance for this estimator involves four individual variance components. Elder *et al.* (1980) provided

$$Var(\bar{z}) = \sigma_x^2/(rn) + [(S-s)/(S-1)][\sigma_w^2/n + n\sigma_a^2(\sigma_x^2 + \sigma_w^2)]/(rs) + \sigma_e^2/(rst) . \quad (4)$$

Also, they established that the expected mean squares (between groups, within groups, and for error) associated with the analysis of variance for the composite sample groups are:

$$EMS_B = (st/n)\sigma_x^2 + [t(S-s)/(S-1)][\sigma_w^2/n + n\sigma_a^2(\sigma_x^2 + \sigma_w^2)] + \sigma_e^2 \quad (5)$$

$$EMS_W = [tS/(S-1)][\sigma_w^2/n + n\sigma_a^2(\sigma_x^2 + \sigma_w^2)] + \sigma_e^2 \quad (6)$$

$$EMS_E = \sigma_e^2 \quad (7)$$

Clearly, the variance of the estimator $\bar{z}$ is equivalent to the value $EMS_B/(rst)$. Thus, the standard error of the estimate of μ_z can be obtained from a standard analysis of variance computation.

References

Brown, G. H. and N. I. Fisher. 1972. Subsampling a mixture of sampled material. *Technometrics* 14(3): 663-668.

Elder, Robert S., William O, Thompson, and Raymond H. Myers. 1980. Properties of composite sampling procedures. *Technometrics* 22(2): 179-186.

Gilbert, Richard O. 1987. *Statistical Methods for Environmental Pollution Monitoring*, Van Nostrand Reinhold Company, New York. 320 pp.

Parrish, R. S., G. O. Ware, C. N. Smith, and P. A. Banks. 1990. Composite sampling techniques for determining pesticide concentrations. *Proceedings of the KSU Conference on Applied Statistics in Agriculture*, Manhattan, KS.

Rhode, Charles A. 1976. Composite sampling. *Biometrics* 32, 273-282.