Battelle
Putting Technology To Work
505 King Avenue
Columbus, Ohio 43201-2693
Telephone (614) 424-6424
Facsimile (614) 424-5263

December 3, 1996

Naval Facilities Engineering
Service Center (NFESC)
Code ESC414
1100 23rd Avenue
Port Hueneme, CA 93043

Attention: Mr. Tim McEntee

Dear Mr. McEntee:

CONTRACT N47408-95-D-0730/0024
Task 6 Review of Statistical Assessment of GW Pathway at Site 09 Allen Harbor Landfill

Please find enclosed my technical comments on the subject report. I have organized my comments into those which I consider most important and less important from a statistical perspective. I hope that you find these comments useful; and I am available to provide further clarification and discussion of the issues if you would like.

Sincerely,

Bruce E Buxton

Bruce E. Buxton, Ph.D.
Program Manager
Statistics and Data
Analysis Systems

BEB:lnl
Enclosures

cc: Mr. Nicholas T. Ta (NFESC)
Mr. Doug Zillmer (NFESC)
Mr. Phil Otis

Review Comments

by

Bruce E. Buxton, Ph.D.
Battelle
December 3, 1996

on

"Statistical Assessment of Groundwater Pathway at Site 09 — Allen Harbor Landfill"

by

NewFields, Inc.
October 31, 1996

Overall I found the report to be well written; however, there are a number of technical questions which I recommend be addressed to more fully document the technical approach that was taken and data analysis results that were obtained. These questions are presented below, separated into those that are more and less important from a statistical perspective.

Major Comments

1. *Page A-7* — The proposed non-parametric correlation analysis which is used to address Conditions 1, 2, and 3 of the null hypothesis (Page A-4) compares data for only a limited number of nearby sampling locations in different media (Tables A-1-1 through A-1-3, and A-1-5). In addition to this "lag 0" type analysis, did the authors consider using a broader cross-correlation analysis that utilizes data paired at all available spatial separation distances? This approach would have at least two potential benefits: (1) it would include many more data pairings in the correlation analysis to more fully search for possible relationships between the contaminant concentrations in two media; and (2) it would eliminate the need to choose only a limited number of pairings of "nearby" sampling locations to be included in the analysis.

2. *Page A-8* — How will statistical significance in the variogram results be quantitatively and objectively assessed; and what is the quantitative decision rule by which the hypotheses related to Conditions 4 and 5 (Page A-5) will be either rejected or accepted? These details are absolutely necessary if the risk pathway analysis is to be "conducted through quantitative statistical hypothesis testing", as stated on Page A-4.

3. *Page A-9* — I agree with the use of directional variograms to help search for potential trends and correlations, as is suggested here; however, I found no such directional variograms in the report. Were these variograms calculated, and if so, what did they show?

4. *Page A-11* — The fact that the analyses for Conditions 1, 2, and 3 failed to detect significant correlations should not be surprising given that the analyses included only 4, 5, or 9 data pairings for each analysis. The fact that no correlations could be detected does not mean that they do not exist. In particular, did the authors attempt to quantify the power of their tests, that is, the probability that correlations of increasing magnitudes could be detected with the limited data pairings that were available?

5. *Page A-12* — By presenting only omni-directional variograms, directional correlation differences can sometimes be "averaged out", leading one to infer that no correlation structure exists. Did the authors calculate directional variograms, and if so, did they show no directional differences so that averaging into omni-directional variograms was an appropriate step to take?

6. *Page A-12* — How was the lack of significance judged in the variograms calculated for Conditions 4 and 5 (see also Comment 2, above)? If significance was qualitatively judged by visual inspection of the variogram plots, then this should be clearly stated along with the plot characteristics that are deemed important (e.g., ratio of apparent nugget variance to apparent sill variance, estimated range of correlation, etc.)

7. *Page A-12* — The logic behind the correlation analysis performed for Condition 6 to assess possible vertical trends should be better developed. I generally think of regression analysis to test for trends. Were data at only two depths available at each sediment sampling location? Did the authors consider other possible tests, such as a two-sample test for the equivalence of mean or median concentrations?

8. *Page A-13* — It appears that the authors use a qualitative, rather than quantitative, assessment to examine Conditions 4 and 5. If this is true, then the overall conclusions must also be stated qualitatively, and the data analysis approach should not be described in terms of a formal statistical hypothesis testing framework. Describing the results in such a framework may tend to exaggerate the statistical significance which the reader may attach to the study findings.

9. *Page A-13* — Data adequacy, in the context of this report, must consider uncertainty related to Kendall's tau statistic (for Conditions 1, 2, 3 and 6) and the variogram statistic (for Conditions 4 and 5). See following comment.

10. *Page A-14* — The assessment of data adequacy based on estimating a binomial proportion seems completely irrelevant for this report's data analysis approach which involves the Kendall's tau and variogram statistics. Can the authors provide a more complete justification for how data adequacy for a binomial proportion relates to data adequacy for Kendall's tau and the variogram?

11. *Page A-15* — Can the authors provide more explanation of the "worst-case situation" that is presented? I interpret "worst-case" to mean that situation where maximum sampling would be

required. Under this interpretation, p=0.5 is the best case, not the worst. For example, if a proportion near p=.20 is to be estimated to within 10% of the true value, then n=400 samples would be required, a number which far exceeds the n=25 samples discussed by the authors.

12. *Figures A-3-1 and A-3-8* — My visual inspection of these two figures suggests that spatial correlation may be evident in the data for these two parameters.

13. *Figures A-3-7, A-3-10 through A-3-15, A-3-17, A-3-18, A-4-1, and A-4-2* — Page A-12 suggests that log transforms of the data were routinely taken to help assess possible spatial correlations with the variogram analysis. Were log transforms taken for these 11 parameters, and if so, how did this affect the variograms?

Other Comments

14. *Page A-4* — The null hypothesis is generally stated along with the decision rule that will be used to reject it. This discussion on Page A-4 does not formally state the decision rules, and it is written in terms of accepting the null hypothesis, not rejecting it.

15. *Page A-5* — Conditions 4, 5, and 6 should be more quantitatively stated; or, if they will be qualitatively assessed (i.e., Conditions 4 and 5) in the data analysis, then they should not be included in the hypothesis testing framework.

16. *Page A-5* — A more detailed explanation is needed for why it is appropriate to assess vertical trends using the non-parametric correlation analysis.

17. *Page A-11* — Have the authors considered how a bias may be introduced to the analyses for Conditions 1, 2, and 3 by routinely including the maximum concentration datum in cases where multi-depth data were available?

18. *Page A-11* — In the third paragraph, should the text "from the Landfill to the Harbor" instead be "from the Harbor to the Landfill"?

19. *Page A-14* — Instead of stating that the data "are devoid of any spatial correlation structure", it may be appropriate to acknowledge that spatial correlation may still be present at scales smaller than the dimensions of the sampling grid.

20. *Page A-15* — Use of EPA's formula for s_p does require lack of spatial correlation in the data, but it is lack of correlation in the indicator data (i.e., indicators of whether the contaminant concentrations are above or below some fixed threshold) that is required, rather than lack of correlation in the concentration data. Assessing potential spatial correlation with indicator variogram analyses is another approach that might be useful for squeezing additional information out of the available concentration data. The indicator variograms are often less affected by extreme outlier data, and so may more readily indicate underlying spatial correlation structures than concentration variograms.

21. *Table A-1-3* — Neither this table nor Figure A-1-2 indicates the locations of samples S9-1, S9-2, S9-3, and S9-4, so it is impossible to judge the appropriateness of the data pairings listed in this table.

22. *Table A-2-1* — I recommend that the sample sizes associated with the estimation of these correlation coefficients be listed along with the significance levels in this table. In this way the reader will be reminded that while statistically significant results could not be demonstrated, the power of these tests to detect correlation is also quite low.

23. *Figures A-3-1 through A-4-2* — It would be useful to confirm the choice of two parameters related to the variogram calculations: (1) the unit lag distance, and (2) the number of lags considered. These parameters are chosen based on the available spatial data coverage (i.e., relative configuration, spacing, and extent of sampling locations), and it is difficult for the reader to determine this coverage from the single Figure A-1-2. I recommend that a couple of additional site sample location figures be included to illustrate the specific data that were included in the various variogram analyses.

24. *Figures A-3-1 through A-4-2* — What are the "distance" units on the variogram figures? Units are also needed for the spatial coordinates listed in Tables A-3-4 and A-3-5.

25. *Figures A-3-2 through A-3-6, and A-3-9* — These figures all show strikingly similar variogram patterns indicating that these six parameters may be highly correlated with each other. Was this cross-correlation checked? If so, the results of the cross-correlation analysis should be briefly discussed somewhere in the report.