75 Pearce Mitchell Place
Stanford, CA 94305
August 24, 1998

John Haggard
General Electric Company
1 Computer Drive South
Albany, NY 12205

RE: HUDSON RIVER

Dear John,

I read the report material which you sent to me relating to Hudson River PCBs, viz., Volume 2C-A Book 1 and Book 2 *Low Resolution Sediment Coring Report* by TAMS *et al.*, dated July, 1998, prepared for the U.S.E.P.A., together with portions of Volume 2C *Data Evaluation and Interpretation Report*, by the same authors, dated February, 1997, containing material related to geostatistical analyses of the 1984 Hudson River PCB data. I also have seen the two reports you sent to me on geostatistical analyses prepared by Davis and Olea for General Electric Company, dated December 14, 1990, and February, 25, 1991 .

In my view, there are a number of important lapses both in the design of the Hudson River PCB study and the analyses of the data that are briefly described in my report accompanying this letter. In short, many of the stated conclusions of the July, 1998, report by TAMS *et al.* are not well supported by the reported statistics. Rather, these conclusions seem to be drawn from interpretations and conjectures that do not have a statistical inferential basis.

Sincerely,

Paul Switzer

COMMENTS ON

1998 U.S.E.P.A
LOW RESOLUTION SEDIMENT CORING REPORT

prepared by

PAUL SWITZER

August, 1998

ISSUES OF DATA SELECTION

1. Since not all core locations from the earlier surveys were to be resampled in 1994, care should be taken that those selected for the 1994 survey are representative of the zones for which PCB inventory estimates are desired. For example, a form of random or systematic selection or areally stratified random selection would remove the possibility of selection bias. Instead, the sampling method used is a poorly documented form of what statisticians call "purposive" sampling – here a misguided attempt to create zones of reduced sediment variability at the expense of unbiased representation of the 1984 sample locations.

2. Seven hot spots were defined by combining certain of the original twelve dredge locations from the 1984 survey. It is not clear whether the decision to combine dredge locations was made before or after the 1994 data were available. If after the data were available, then another source of selection bias is created because combining data can be used to purposively exaggerate statistical significance. This effect can be seen, for example, by comparing Figure 4-21 with Figure 4-22.

3. Not all of the 1994 data that were collected were used in the statistical estimates. An important issue is the removal of nearly 40% of the data because of potential cross-contamination from higher to lower segments of a core. The severe data rejection rule conforms to *a priori* ideas of what a core sequence should look like rather than relying on explicit evidence of cross-contamination. Because the fraction of data removed from the analysis is large, the susceptibility to selection bias is also large. Calculations showing that some statistics are not seriously affected by this form of data screening do not adequately address this issue.

4. It was reported that the lowest concentration data were excluded from "many of the subsequent analyses" [page 2-16] so that the remaining data would look like a lognormal distribution. The fact that the low concentration data are hard to quantify does not change the fact that they are nevertheless informative. Trimming data to achieve a desired distributional shape violates important statistical principles related to selection bias.

5. As an example of how purposive data selection can create an impression of a statistical relationship compare the weak associations in Figure 3-2 with the strong associations indicated by the selected data subset of Figure 3-8. This example shows clearly the perils of censoring data rather than modeling data.

CALIBRATION ISSUES

1. The SSW correction [see Figure 4-17 and Table 4-3] applied to the 1984 data is itself based on the PCB measurements. The net effect is to introduce an implicit non-linear rescaling into the PCB data which is difficult to analyze. This dry soil weight SSW factor should be based on some physical parameter other than PCB itself. Furthermore, stronger evidence is needed to demonstrate that such a recalibration of the 1976-78 data makes them comparable to the 1994 data.

2. The measurement methods for the 1976-78 survey, the 1984 survey, and the 1994 survey are different [see comments on page 4-21, for example]. However, there are no cross-calibration data presented which analyze the same samples by both methods. By implication, the subsequent statistical analyses assumed that relative calibration errors were strictly zero. How could the 1994 study have been undertaken without such calibration information?

3. The extrapolation of grab samples to make them comparable to 12-inch core segments is not well documented in the report. The "factor" used for extrapolation appears to be something like the mean PCB ratio between the top four inches and the top 12 inches seen in core samples [page 4-21]. Of course, such extrapolation introduces additional error beyond the measurement error associated with core sampling but the extrapolation error is not accounted for in the analysis.

THE REGRESSION FALLACY

Surprisingly, the report does not recognize the "regression" fallacy. The report makes much of the observation that lower 1984 concentration locations tended to be followed by later increases while higher 1984 concentrations tend to be followed by decreases. This kind of analysis is found throughout the report without any recognition of the "regression-toward-the-mean" effect: for *any* collection of paired measurements which are positively correlated, which these data emphatically are, higher initial concentrations will tend to show a decrease on average and lower initial concentrations will tend to show an increase on average -- *in the absence of any distributional changes*! [See Fig 4-11 to Fig 4-16 for example, as well as much of Section 4 and statements in the Executive Summary]. This fallacy may be understood by simply supposing that the true concentrations were *exactly* the same in 1984 and 1994 and that reported concentration differences were due entirely to analytic variability. Even in this case, splitting the data into two parts as the authors have done would demonstrate that higher 1984 concentrations tend to be followed by lower 1994 concentrations and vice-versa.

The geostatistical evidence from the 1984 survey shows that very short-scale spatial variability is often comparable to total variability [see Figures 4-9 to 4-12 and 4-14 to 4-16 in the 1997 report]. This fact, by itself, can be used to show that even small location errors induce a large spurious regression effect that would account for most of the reported 1994 PCB decrease at locations with high 1984 PCB.

ISSUES RELATED TO STATISTICAL CALCULATIONS

1. The MVUE method for estimation of means requires an exact lognormal model. Otherwise the method is biased. For example, the MVUE is sensitive to procedures for handling low concentration data. Censoring and BDL data create special problems for the MVUE method. The straightforward sample average is less sensitive to prior assumptions about the distribution of the data.

2. The lognormal unbiasedness correction was inconsistently applied. For example, compare formulas on page 4-28 and in Table 4-13. In the latter case, the supposed MVUE ["minimum variance unbiased estimate"] for the near-shore sediments is neither unbiased nor minimum variance, even if the data were sampled from a precisely lognormal population distribution.

3. While data were collected on covariate parameters such as grain size, the estimates of PCB inventory change did not incorporate the covariate information in any statistical model. It is not clear where or how the relationship between fine grain sediments and elevated PCB was used to improve estimates of PCB inventory [page 3-34].

4. Simply reporting correlations between individual covariate parameters and PCB measures, without regard to the multivariate nature of the data, does not take advantage of the potentially more powerful but widely used multivariate statistical methods such as partial correlation or multiple regression analysis. See the discussion on page 3-15 as an example. Failure to examine the multivariate statistics can also lead to mistaken interpretations such as that found on 3-19 "The correlation of ΔMV and MDPR values is largely a ramification of the correlation with total PCB", as well as a similar confused statement about correlations on page 3-25.

5. Hot spot PCB quantity differences between the 1976-78 and 1994 surveys are assessed by the ratio Δ_t described on page 4-40. The explicit claim is that differences "of more than a factor of two are considered significant and likely to be beyond the level of uncertainty". I could find absolutely no statistical basis for this claim and no supporting analysis!

6. Upper 95% upper confidence limits were used to compare sediments in Table 4-13. Differences among these upper limits may be due to mean differences, variability differences, or differences in the number of samples, or some combination of these. It is misleading to make such comparisons.

10.1786

7. The report contains no statistical analysis of variance components that could be used to make better judgements regarding the significance of observed differences. Some examples of large random components of variability include the following. On Page 2-18 we read that split-pair samples differed by 36% on average. Table A-3 suggests that field replicate-pair congener analyses differ by more than 20% about half the time. Geostatistical studies of the 1984 survey report that "it was not unusual for samples taken only a few meters apart horizontally to exhibit order-of-magnitude differences in PCB concentrations" [page 4-25, 1997 report]. The reported variograms indicate that very short-scale spatial variability in sediment PCB is often comparable to the total variability observed between different sample locations. Without adequate accounting for random error components, it is hard to appreciate reported differences between the earlier and later PCB surveys.

8. Statistical analyses of the Δ_M ratio measure of PCB molar inventory change are based on adding the arbitrary constant 2.0 to all ratios and then taking the logarithm. It is not reported how sensitive are the analyses to this particular fudge factor 2.0 . However, it is likely that the low end of the distribution is noticeably affected by this added factor, thereby affecting statistics in the context of a lognormal model.

9. It appears that zonal PCB inventory estimates were obtained simply by multiplying averages of core-based concentrations [g/m^2] by estimates of zonal areas. These estimates take no account of the spatial configuration of the cores within zones and make no use of the geostatistical considerations. Although the 1994 data are not sufficient to develop the appropriate geostatistics, the geostatistical models based on the more extensive 1984 survey could have and should have been used.

10. Estimates of PCB inventory, inventory changes, and dechlorination losses [see Table 4-8 as an example] are not accompanied by statistical estimates of uncertainty. Good statistical practice dictates the quantification of uncertainty – first as a measure of the adequacy of the underlying information and second as a tool in rational decision making. The uncertainty associated with inventory estimates derives from sampling variability, measurement error, and spatial interpolation errors, all of which propagate to the final inventory estimates.

INCONSISTENCIES

1. The ^{7}Be data sometimes did, and sometimes did not show association with PCB levels or imputed scour of PCB, depending on how the data were divided or otherwise manipulated. See pages 3-23 and 4-38 for example. Nevertheless, the Executive Summary states categorically that "beryllium-7 was shown to be a statistically significant indicator of inventory loss" [ES-3]. For hot spots the report states that " [^{7}Be] evidence for recent deposition apparently contradicts the strong PCB profile evidence for long-term loss sediment loss or lack of burial" [page 4-35].

2. It is reported that the molar inventory changes between the 1984 and 1994 surveys [page 4-14] indicate no statistically detectable trend. However, the report also claims that a relative measure of molar change, Δ_M , shows statistically detectable decreases, probably reflecting the regression fallacy described earlier. In this connection, the report makes the inaccurate statement on page 4-15 that for relative measures of change "much of the analytic variability can be diminished in importance relative to real change".

3. The high-resolution 1984 survey did not confirm that the amount of implied dechlorination is proportional to total PCB or that it followed the same dechlorination pattern as the low-resolution 1994 data [pages 3-5, 3-9, 3-35].

CONCLUSIONS VERSUS DATA

1. Statistical analyses were not used to support frequent statements regarding lack of burial. Indeed, high resolution core data [Figure 4-24 is an example] indicate that the concentration peak can rise and fall sharply within a 9" or 12" segment, representing one or more decades of deposition. The fact that the top segment of a low-resolution core shows higher concentration than the succeeding segment does not, by itself, imply that the peak concentration is at the surface and that burial is absent. On page 3-10 the report itself states that "it is likely that a low-resolution core segment would span a range of two or more orders of magnitude...over a nine inch interval".

2. The simple "sign test" shows that the frequency of locations with PCB increases versus those with decreases is not statistically different from 50%. There are many examples, such as the comparisons in Figures 4-4 and 4-21. The point is that the limited size of the 1994 survey makes the results look much like the results of coin tossing. The non-statistical rationalizations in the report, however salient, should not be confused with statistical inference.

3. Many far-reaching statements in the report are not actually derived from a statistical analysis of the data, even though such non-statistical hypotheses and conjectures are intertwined with statistical reporting. Examples of such are statements referring to scour, resuspension or redistribution of sediments, dechlorination limits, underestimation of prior inventory, probable losses due to incomplete cores, and the relation of Aroclors to PCB concentrations.

December, 1997

PAUL SWITZER
CURRICULUM VITAE

VITAL DATA

Birth Date: March 4, 1939
Birth Place: St. Boniface, Manitoba, Canada
Citizenships: U.S., Canada

EDUCATION

University of Manitoba, B.A. (hons.), Mathematics, 1961;
Harvard University, A.M., Statistics, 1963;
Harvard University, Ph.D., Statistics, 1965.

POSITIONS

Professor, Statistics Department, Stanford University, 1965-
Professor, School of Earth Sciences, Stanford University, 1965-
Visiting Professor, Singapore National University, 1997
Visiting Professor, University of Rome La Sapienza, Rome, 1992
Visiting Scientist, Goddard Institute for Space Studies, New York, 1991
Visiting Professor, Ecole des Mines, Fontainebleau, France, 1977 & 1988
Visiting Professor, Massachusetts Institute of Technology, 1984
Visiting Professor, E.T.H., Zurich, Switzerland, 1984
Visiting Scientist, C.S.I.R.O. Division of Mineral Physics and
Division of Mathematics and Statistics, Sydney, Australia, 1983
Visiting Fellow, Japan Society for Promotion of Science, Tokyo, 1983
Chairman, Statistics Department, Stanford University, 1979-82
Visiting Scientist, Environmental Protection Agency, Washington, 1981 & 1991
Visiting Professor, Statistics Department, Princeton University, 1974
Visiting Professor, Imperial College, University of London, 1973

PROFESSIONAL ORGANIZATIONS

Institute of Mathematical Statistics
American Statistical Association
International Statistical Institute
Royal Statistical Society
International Association for Mathematical Geology
American Geophysical Union
Canadian Statistical Society

EDITORSHIPS

Editor, Statistical Science, 1995-1997
Editorial Board, Chapman & Hall series "Interdisciplinary Statistics", 1992-95
Consulting Editor, Canadian Journal of Statistics, 1992-94
Editorial Advisory Board, Atmosphere Environment, 1992-
Editorial Board, Oxford University Press Series on Spatial Information Systems, 1989-
Editor, Journal of the American Statistical Association, 1986-88
Editor, Statistical Science, 1984
Statistics Editor, Littlefield, Adams Publishing Co., 1983-88
Associate Editor, Canadian Journal of Statistics, 1980-85
Associate Editor, Journal of the American Statistical Association 1975-78
Associate Editor, Mathematical Geology, 1974-76

HONORS AND AWARDS

Presidential Invited Speaker, Canadian Statistical Society, 1996
Nomination, Phi Beta Kappa Excellence in Teaching award, 1996
Topical Lecturer, American Association for the Advancement of Science, 1994
Distinguished Achievement Medal, American Statistical Association,
Section on the Environment 1993
Presidential Neyman Lecturer, International Statistical Institute, 1991
Closing Plenary Lecturer, European Meeting of Statisticians, 1991
Fellow, International Statistical Institute, 1982
Fellow, American Statistical Association, 1980
Fellow, Institute of Mathematical Statistics, 1980
Best Paper Award, International Association for Mathematical Geology, 1980

PUBLIC AND PROFESSIONAL SERVICE [since 1990, selected]

Research Proposal Reviews: NSF, Netherlands NSF, NSERC Canada, 1997
Promotion Review, Dept of Statistics, University of Washington, 1997
Program Review, Dept of Statistics & Operations Research, Kuwait University, 1996
Promotion Reviews, Stockholm University, Penn State University, 1996
National Advisory Council on Environmental Policy and Technology,
    US Environmental Protection Agency, 1995-
Environmental Statistics Technical Advisory Committee,
    American Statistical Association, 1995-97
Board of Directors, Societal Institute of the Mathematical Sciences, 1995-
Committee on Global & Environmental Change, American Geophysical Union, 1993-
Student Conduct Legislative Council, Stanford University, 1993-95
Humanities & Sciences Council, Stanford University, 1993-94
Statistics Steering Committee, National Center for Atmospheric Research, 1993-
Promotion Reviews, University of Minnesota, Woods Hole Oceanographic Inst,
    UC Santa Barbara, New York University, 1995
Promotion Reviews, New York University, Northeastern University,
    University of Wisconsin, 1994
Promotion Reviews, University of Chicago, NYU, Washington, Arizona,
    UC Berkeley, Virginia, 1993
Research Proposal Reviews, National Science Foundation,
    SERC [UK], IMA, NOAA, others [continuing]
Editorial Reviews, Urban Atmosph., Oxford Univ Press, IEEE,
    Math Geol, EPRI, others
Organizer, Session on Climate Patterns and Change, ASA Annual Meeting, 1995
Organizer, Session on Data Analysis for Physical Sciences, IMS Annual Meeting, 1994
Organizer, Session on Global Climate Change, Inst.Math.Statistics Annual Meeting, 1992
Scientific Committee, Fifth World Geostatistics Congress, 1992
Organizer, Workshop on Monte Carlo Methodologies in Human Exposure Analysis,
    EPA 1992
Panel for Earth Sciences & Statistics, National Science Foundation, 1991
Program Committee, Institute of Mathematical Statistics, 1991-92
Organizer, 5th Internat. Meteorology & Statistics Congress, 1991-92
Review Panel, Mathematical Sciences Division Office of Naval Research, 1991
Climate Science Advisory Committee, EPRI, 1990-
Committee for Statistics in the Physical Sciences, Bernoulli Society, 1990-94
Organizer, AAAS Annual Meeting, Earthquake Prediction and Validation, 1990-91
Board Member, American Statistical Association, 1990
Organizer, Institute for Mathematics Applications,
    Summer Program for Environment, 1989-92
Chair, Review Committee on Undergraduate Mathematical Sciences, Stanford Univ., 1989
Program Committee, 2nd World Congress Bernoulli Society, Upsalla, 1989-90

SELECTED PUBLICATIONS [after 1990]

"Spatial Covariance Estimation for Monitoring Data" with C. Loader,
Statistics in the Environmental and Earth Sciences, 1991.

"Undiscovered Oil and Gas Resources:
An Evaluation of the Department of the Interior's 1989 Assessment Procedures"
[NRC panel], National Academy Press, 192pp, 1991.

"Spatial and Temporal Statistics for Environmental Problems",
Bull.Int.Statist.Inst. 1992 pp 315-328.

"Derivation of an Indoor Air Averaging Time Model from the Mass Balance Equation"
with W. Ott, J. of Exposure Analysis and Environmental Epidemiology, 1993, pp 113-135.

"Carbon Monoxide Exposures Inside an Automobile Traveling on an Urban Arterial Highway"
with W. Ott & N. Willits, J. Air & Waste Manage. Assoc. 1994 pp 1010-1018.

"Spatial Interpolation Errors for Monitoring Data" with G. Host & H. Omre,
J. Amer.Statist.Assn, 1995 pp 853-861.

"Time Trend Estimation for a Geogrpahic Region" with K. Solna,
J. Amer.Statis.Assn, 1996 pp 577-589.

"Ambient Sulfur Concentrations Near Grand Canyon as a Function of Fluctuating Loads
at the Mohave Power Project", with L. Enger et al.
Atmospheric Environment, 1996 pp 2552-2564.

"Differential Exposure Misclassification in Case-Control Studies
of Environmental TobaccoSmoke and Lung Cancer" with M. Levois,
Journal of Clinical Epidemiology 1998.

General Electric Company
Corporate Environmental Programs
One Computer Drive South
Albany, NY 12205
Telephone: (518) 458-6618
Fax: (518) 458-1014

EPA's Key Hudson River Questions With Answers Provided by QEA's Upper Hudson Model

Q: When will PCB levels in fish populations recover to levels meeting human health and ecological risk criteria under continued no action?

A: Under a no action or natural recovery scenario, average PCB levels in largemouth bass will reach the FDA's 2 ppm threshold by about 2000 in the largest part of the Upper River (the 29 miles between Schuylerville and Troy) and by about 2010 in 11 miles from Fort Edward to Schuylerville.

Q: Can remedies other than no action significantly shorten the time required to achieve acceptable risk criteria?

A: Elimination of the residual PCB sources associated with the area of GE's Hudson Falls plant site accelerates the decline of PCB levels in fish to the 2 ppm threshold in the Thompson Island Pool area. If the upstream source were eliminated now, average PCB levels in largemouth bass from Fort Edward to Schuylerville would reach the FDA threshold in about 2006, four years sooner than under no action. Average PCB levels in largemouth bass in the Schuylerville to Troy area would still reach the FDA limit in about 2000.

Dredging the historic PCB hot spots from Fort Edward to Schuylerville would not accelerate the decline of PCB concentrations in fish beyond that which would occur with no action. Under both scenarios, average PCB levels in largemouth bass would decline to 2 ppm by 2000 in the area of Schuylerville to Troy and by 2010 from Fort Edward to Schuylerville.

Dredging all of the fine or cohesive sediments in the area between Fort Edward and Schuylerville would not accelerate the decline in PCB concentrations in fish from Schuylerville to Troy; they would still reach the FDA threshold in about 2000. In the area of Fort Edward to Schuylerville, average PCB levels in largemouth bass would reach the FDA limit in about 2009.

Estimated Time for Upper Hudson Fish to Achieve 2 ppm FDA Safety Threshold*

Based on Quantitative Environmental Analysis/GE Upper Hudson Model

Clean-up Scenarios	Northumberland/Schuylerville to Troy *(29 Miles)*	Fort Edward to Northumberland/Schuylerville *(11 Miles)*
Baseline Condition *Natural recovery; no additional reduction in PCBs coming from Hudson Falls area*	**2000** (2004)	**2010** (2014)
Upstream Source Control *Elimination of residual PCB sources in Hudson Falls area through ongoing GE clean-up project*	**2000** (2002)	**2006** (2011)
Dredging		
• *Hot spots*	**N/A****	**2010** (2012)
• *All near-shore fine sediment*	**N/A****	**2009** (2011)

Bold numbers indicate best estimate. Shaded numbers in parentheses indicate reasonable worst-case scenarios, e.g. slower than historical sedimentation rates.

*** Based on average PCB concentrations.**

**** Through the river's natural recovery process, average PCB levels in Upper Hudson fish will reach the FDA threshold before EPA's June 2001 deadline for a decision on the Upper Hudson or before dredging could begin.**