23 Waterview Drive
Saratoga Spgs.,NY
12866-8791
Dec. 23, 1999

Mr. Douglas Tomchuk
USEPA-Region 2
290 Broadway
New York, NY 10007-1866

Dear Mr. Tomchuk:

I recently changed my residence and belatedly received the December letter advising of submissions to peer review. I would hope there is still time to include my comments below in the next submission of charges to the peer group.

At this point in the reevaluation process and agenda my concern is with proper hypothesis testing, or rather the lack of same, in respect to the Hudson River model(s) that has been constructed, and the estimates derived therefrom - which now are proffered as proven conclusions without as yet any analysis of error or test of validity. Currently, two multivariate function models have been advanced (TAMS/EPA and QEA/GE) to explain Hudson River processes and response, and used to generate estimates or predictions of future conditions or state of the system under several different scenarios.

The models are hypotheses and their products are estimates of outcomes, or predictions, and must be formally tested for significance or validity by accepted methods before they can be considered meaningful conclusions. A model can be complex and claim to have integrated all relevant variables, and still be very wrong as any survey of science demonstrates.

In brief, hypothesis testing of multivariate functions involves an estimation of the variance of each separate variable (such as PCB concentration water at a specific time and place), a pooling of all variances according to established rules to estimate total variance, and then making a comparison to the appropriate probability distributions of differences to determine the likelihood that the hypothesis is true (i.e. provides a valid description of the facts and how they are related).

Predictions from the hypothesis (model) are then customarily reported with reference to the range of error that can be attached to a particular result for a given probability, commonly expressed as confidence limits at 90 to 95% certainty. The larger the range of values (or error limits), the weaker and less efficient the hypothesis as a descriptive and predictive tool.

Both models predict that PCB loading in the system and concentrations in fish will decrease with time; they differ, however, in predictions of the rate of decrease or the time to reach a fixed reference value or state. Unless the error limits are so large that they mutually overlap (and render distinction moot), both models cannot be "correct".

Considering the implications of the differing model predictions for the need of remediation, its practical implementation, and ultimate cost/benefit vs. risk basis, a rigorous multivariate statistical testing procedure should be employed at once. This is not, by any means, a simple matter of just calculating a few standard deviations and means from the data, and then using some "standard statistical calculations" to derive confidence limits, - as some members of the Science and Technical Committee may feel - because none of the assumptions inherent in using any short-cut formulation are satisfied in the present case.

One of the major difficulties now in evaluating either existing model is not the lack of a determination of confidence limits, but far more basic in the lack of any systematic reporting, or assessments, of variance in the data used for model calibration and congruity or "look-back" comparisons.

For one example, suspended sediment loading is a model input variable. Its variance includes contributions from horizontal and vertical position of the sample, River mile, flow or discharge rate, time of sampling relative to a discharge event cycle, and relative to prior events; time of year, proportion of sediment that is external in origin and not resuspended, and lastly the sampling method itself. An initial analysis of variance study can optimize the sampling method (not done) to reduce the variance contributions from some of the above, but not others. A final contribution to variance is provided by the multiple sample points or stations used.

The resultant total variance above is integrated into the model functions, and becomes a part of any output where suspended sediment concentrations are used as an input variable (example: depth of scour in 100yr flood event; downstream transport of PCBs). It should also be noted that the mere estimation of a sample variance says nothing about the validity of any assumed relationships among the input variables; each relationship is a smaller scale hypothesis in itself that must be tested for significance before incorporation in the model.

The TAMS/EPA model, for example, assumes that discharge rate and suspended sediment concentrations are simply related (and hence a relation to scour); aside from the lack of variance determination, this assumption has not been tested by regression or any other means in any documentation I have received. This is just one example, many others can be cited.

In short, I request that the responsibility for testing both Hudson River models by appropriate statistical analysis procedures, for internal coherence and significance level of the predictions made, be charged to peer review; further, a requirement that results from the model of higher significance rank be incorporated as definitive in the reevaluation investigation be considered.

I also have some items concerning the data base and documentation used for the Human Health Risk Assessment that I would like submitted to peer review, which I assume will be in a subsequent session.

As a scientist, the requested charges are nothing more than the standards and requirements of scientific proof which I am subject to by peer review of research and publication in my own profession, and I expect nothing less of any endeavor that claims to furnish the same.

Please acknowledge as to whether my request is included in the EPA charge submitted to peer review, and in what form if not verbatim.

Thank you for your attention and consideration in this matter,

Very truly yours,

George W. Putman, PhD

cc: J.Haggard, GE
G.Hodgson, SCEMC
J.Davis, NYSDOL