3/15/2005

External Peer Review of Scientific Portion of Proposed California Ocean Plan Amendment.

Reviewer: George Sugihara
Professor
Scripps Institution of Oceanography
University of California San Diego
La Jolla, CA 92093-0202
Tel: (858) 534-5582
Mobile: (858) 829-8111
gsugihara@ucsd.edu

The scientific portion of the proposed regulatory plan concerns a suite of statistical methodologies to determine if an effluent discharge has a "reasonable potential" to exceed a Table B water quality objective. Each methodology is suggested to be more appropriate in each specific data environment (i.e., large sample sizes of uncensored monitoring data, truncated data sets, and data sets of very small sample sizes). The main idea for combining these methods into a coherent scientific protocol is contained in the flow chart Figure VI-1. This chart gives an algorithmic recipe of contingent actions/analyses to be followed depending on each data environment.

While I have some questions and personal concerns about how these methods may be applied (as described in detail below), I can say that according to statute mandate for scientific peer review, I believe these methods, as they stand as narrowly defined statistical questions, are scientifically sound and reasonably represent the state of the art. I believe the composite method proposed here by the Ocean Unit staff is superior to the existing EPA water quality protocol (the TSD procedure). I would encourage that these methods be adopted promptly.

The statistical methods proposed by the Ocean Unit staff, which I shall now discuss in detail are as follows:

1. The use of normal tolerance factors and the lognormal distribution assumption when calculating an upper one-sided confidence bound on the 95th percentile with the parametric approach (UCBL).
2. The use of robust regression on order statistics when data are not too severely truncated (Helsel and Cohn 1988).
3. The binomial approach for comparing severely truncated or sparse data sets with a regulatory standard.

Specific Comments:
(i) The UCBL method advocated here involves calculating an upper one-sided confidence bound for a completely unveiled distribution of values, as opposed to the TSD procedure which involves setting an acceptance threshold as a multiple (k) relative to a maximum

observed value (X). Both methods involve the underlying assumption of lognormality, however it appears that TSD is inferior to UCBL in statistical power in that it depends critically on a single observed maximum value. The expected maximum value X in TSB obviously increases with sample size, and as a single value it will be less reliable as a sample statistic (in terms of higher variance in the estimates) in small sample sizes than UCBL (computes summary statistics). Indeed, in very small sample sizes TSB advises the use of a somewhat arbitrary recipe for computing the multiplicative TSB factor, k. I think UCBL is superior to TSD at all sample sizes.

(ii) In general, with moderately small sample sizes, TSD provides a less conservative estimate except when large outliers are detected. However, as I understand it, provided such extreme (fat-tailed) events are not measurement artifacts, they are probably important for the policy objectives, and should not be averaged out as in UCBL, but should probably be taken into account. This is apparently the rationale behind the GSL protocol, and I am pleased to see that this was explicitly taken into account in the decision tree protocol of Figure VI-1 (see comments below on nonparametric tests).

(iii) It seems like it might be useful to have some specific guideline as to the precise sample sizes required to invoke specific procedures. General guidelines are given, and perhaps it is problematic (or even misleading) to be more precise.

(iv) I was pleased that the authors showed that the UCBL is robust to several different underlying distributions (all of which have the same skew). I think the ones chosen are adequate and appropriate given the generic property that the data are generally skewed and look lognormal, however it might be nice window dressing if these distributions represented extremes that imply generality in the way that Tukey has suggested (triangular, U-shaped and uniform). Again this is just a cosmetic suggestion, and I am happy with the current spectrum of distributions chosen to verify robustness as they are close to the ones that are observed.

(v) This next comment is not intended as a fair criticism of UCBL. It goes beyond the current perspective of water quality standards, and should be thought of as grist for future research. The rationale for genesis of the observed lognormal distribution of effluent concentration values has to do with sequential dilutions. This appears similar in motivation to sequential breakage (Sugihara et. al. 2003. PNAS) with each dilution event effectively acting like a multiplier ($e_t < 1$) that is independent of effluent concentration at time t. It is the same rationale that Kolmogorff invoked to explain lognormality in particle size distributions resulting from sequential breakages. Such mechanisms suggest an interesting canonical coupling between the mean and variance of the resulting lognormal (Seigel and Sugihara 1982 J.Appl. Stat). This would mean that the variance in effluent concentration might be sensitive to distance from source or to the complexity of the cascade of dilution events. If this is true, then it seems likely that the interpretation given to 95% confidence limits or indeed any statistical procedure involving first and second moments might be modified to take this into account. Again, I see this as a second order problem that would be interesting to pursue as a future research project that may or may not change the UCBL procedure. UCBL is clearly better than anything else currently in place.

(vi) The use of Helsel and Cohn's (1988) robust regression on order statistics when data are not too severely truncated appears straightforward. I am pleased that Clifford Cohen's early work in Technometrica is getting attention in this problem (I have an stack of old computer cards from graduate school with his estimation algorithm written in Fortran!). I commend the authors for recognizing the importance of data truncation in estimating the parameters for the underlying distribution.

(vii) The problem of small sample sizes (especially with truncation) remains my major concern, however, I believe that the simple binomial test as described is a reasonable and technically sound thing to try. The bottom line however, is I would not place a bet on it. There is large unavoidable uncertainty in any statistics that one tries to do with small sample sizes. Thus, for example, although the statistical rationale is clear, from a policy objective point of view I remain uncomfortable (eg. with the 80%, n=16 recommendation). Granted, a more stringent recommendation might not be feasible. This is not a technical criticism so much as a logical one.

(viii) Table 3 overstates it's case in that it suggests that UCBL gives a (credible) estimate of effluent variability at ALL sample sizes. (not n=1 at least).

Final Global Comment:
Again, seeing statistics as a tool for achieving more rationality in decisions, it seems that more attention should be focused on the protocol for data collection. I am sure you agree that conclusions from particular statistical procedures are only as powerful as the quality of the raw data to begin with. Insuring there is an accurate (low measurement error) sufficient (n-large) sample of some putative universe of values is key. This echoes comment (iii) above. In the case of highly censored data, if possible it would make sense to collect it in such a way that there is less truncation (eg, closer to the source, or with a more sensitive assay). Statistical creations can amplify the meaningful signal in these data, but ultimately when the data are excessively meager or unreliable they might promise far more than they can deliver. Nevertheless, I would encourage rapid adoption of the methods proposed here, as a clear step in the right direction.

UNIVERSITY OF CALIFORNIA, SAN DIEGO

UCSD

BERKELEY • DAVIS • IRVINE • LOS ANGELES • RIVERSIDE • SAN DIEGO • SAN FRANCISCO

SANTA BARBARA • SANTA CRUZ

GEORGE SUGIHARA
John Dove Isaacs Professor
of Natural Philosophy

Scripps Institution of Oceanography
and California Space Institute
9500 Gilman Drive, 0202
La Jolla, California 92093-0202
(619) 534-5582
(619) 534-1711 FAX
sugihara@ucsd.bitnet

Dear Steve,

Here is my signed review as promised.

Sincerely,

George Sugihara