Correspondence

Comment on "Method Detection Limits in Solid Waste Analysis"

SIR: In the recent paper by Kimbrough and Wakakuwa (*1*), they concluded that the U.S. EPA's method detection limit (MDL) procedure [40CFR136b, (*2*)] "produces accurate and precise results only in interference-free conditions", and reported on an interlaboratory study which indicated that "approximately two-thirds of the reported MDLs produced significant errors". While we believe the conclusions are valid and suspect that a large fraction of MDLs estimated according to the U.S. EPA procedure have unacceptably large errors, we take issue with the authors on some points: (1) detection definitions; (2) accuracy; and (3) quantitation.

(1) Detection Definitions. The authors concoct new and unwise definitions of the terms "false positive" (FP) and "false negative" (FN). These definitions are intended to reflect the "data users' perspective", but unfortunately they turn the discussion of appropriate limits inside out. The traditional definition of a FP is *detecting* an analyte in a particular matrix when, in fact, none is present—above a possible background level (and also allowing for possible matrix interferences). Kimbrough and Wakakuwa redefine a FP as detecting an analyte above a reporting limit (however chosen) when, in fact, the concentration is below the limit. Similarly, the traditional definition of a FN is *not detecting* an analyte when, in fact, it is present at some particular concentration. The FN rate would thus be a (declining) function, FN(conc), of the true concentration. The particular FN rate of interest for environmental MDL contexts is the rate of nondetection at the MDL, FN(MDL). Kimbrough and Wakakuwa redefine a false negative as reporting a "less than" value less than the true concentration.

As Kimbrough and Wakakuwa admit in the paper, these new definitions have "one problem", they ignore measurement uncertainty. In trace level measurement (environmental or otherwise), measurement uncertainty *cannot be ignored.* Instead, it is essential to metrologically defining the MDL; an MDL to be used for reporting or regulatory purposes should be constructed to have good true positive and true negative detection properties—which protect the interests of the regulator and the regulated party. The MDL procedure provided in 40CFR136b nominally provides a FP rate of 1%, but since it does not account for calibration error, the actual FP rate is commonly 2% or higher (*3*). Additionally, as emphasized by Kimbrough and Wakakuwa, the FP rate can be driven higher still since 40CFR136b inadequately accounts for matrix background. The FN rate of 40CFR136b is not deliberately set by the procedure, but has been shown to be about 50%. This means that a regulated party may be marginally above a compliance limit but will have only a 50:50 chance of detecting the violation with a monitoring measurement. Kimbrough and Wakakuwa find this unacceptable, and we entirely concur.

Unfortunately, Kimbrough and Wakakuwa's new detection definitions do not help because they are tied to how

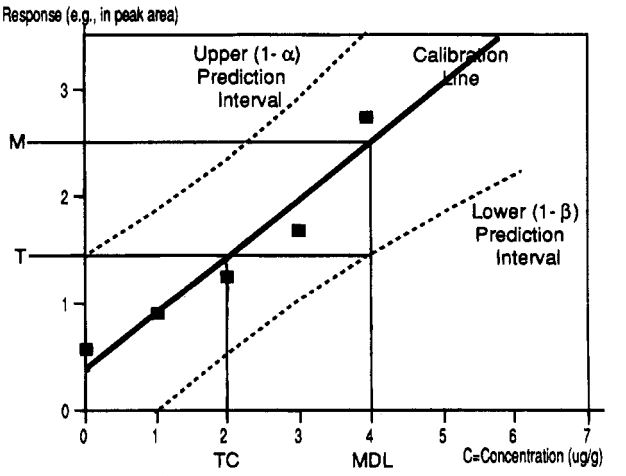

FIGURE 1. Trace-level calibration and measurement error.

results are reported rather than detection probabilities at different concentrations. By their definition of FN, any regulatory limit used in a "less than" reporting format will have a 50% FN rate. As they state, half of the measurements produced by unbiased measurement of a sample spiked at 5 μg/g cadmium can be expected to be less than 5, so a 50% FN rate will be obtained if those measurements are reported as "less than 5". Once again, this is clearly unacceptable, but in this case it is an artifact of Kimbrough and Wakakuwa's definition of FN.

Hubaux and Vos, in their seminal article in *Anal. Chem.* (*4*), presented a way to compute a MDL (actually, a MDL and response threshold pair) which provides specified levels of α = FP and β = FN detection probabilities. As Figure 1 illustrates, the response threshold for detection, T, is selected to ensure that the probability of a blank response exceeding T is α, and the MDL is the concentration at which the probability of a response exceeding T is $1 - \beta$ (it has the expected response, M). For detection at or below the MDL, probabilities are all that matter, not quantitation. The Hubaux–Vos MDL can be shown to be the lowest concentration with a relative measurement error (RME) below ±50% at confidence $100 \times (1 - \alpha - \beta)$. Equivalently, the MDL is the lowest concentration with >0 significant digits (at some confidence level and allowing for fractions of a digit in base 10 arithmetic). This is a reasonable definition of trace-level detection, historically used throughout signal processing theory and application: the point at which there is *sufficient* information content to make a binary decision at acceptable risk, but no more (*5*). It suggests a statistically sound approach to reporting response R (with concentration C) is as shown in Table 1.

(2) Accuracy. This is the primary area in which Kimbrough and Wakakuwa criticize the 40CFR136b procedure. They note that 40CFR136b is based entirely on measured variance, but correctly assert there are two sources of bias which should not be ignored: laboratory and matrix. While it is true that the EPA's procedure ignores lab bias, it does *recognize* the possibility of matrix interfer-

0013-936X/95/0929-0279$09.00/0

TABLE 1

Three Possible Outcomes of Trace-Level Testing/Measurement, Using Hubaux–Vos's Calibration-Based MDL with Response = R, Measured Concentration = C, FP Rate = α, and FN Rate = β

result	sample case (Figure 1)	what should be reported	rationale/interpretation
$R < T$	$R = 1$	ND (for T, MDL)	probability, P_0, that a blank sample would produce a response at this level or higher is greater than α
$T < R < M$	$R = 2$	detect but do not publish C (for T, MDL)	$P_0 < \alpha$, but the relative error in measurement $> \pm 50\%$, so there is not even a fraction of a significant digit
$M < R$	$R = 3$	detect and publish C (for T, MDL)	$P_0 < \alpha$, and there is at least a fraction of a significant digit; also, probability of detection at this concentration $\geq 1 - \beta$

ence (6). Unfortunately, the variation introduced by blank subtraction is not statistically incorporated in the MDL. Note that the Hubaux–Vos procedure based on non-zero intercept calibration using the matrix of interest *does* incorporate matrix interference in the MDL by way of the prediction error interval. Interlaboratory bias modeling is intrinsically more difficult.

(3) Quantitation. Kimbrough and Wakakuwa confuse detection and quantitation. In their introduction, they state "A less than value on an analytical report is as much a data point as a numerical value and should be determined with equal precision and accuracy". This is misguided, and the distinction is critical. Detection is like testing. It is done at the lowest limits of a usable signal to noise ratio. Its result is binary; providing information with the lowest possible content. Quantitative measurements have more content and are far more valuable to scientists and engineers, but are made at concentrations with higher signal to noise, well above the MDL. For quantitation, the Hubaux–Vos approach has been extended to compute a PQL defined as the lowest concentration at which measurement with one significant digit can be made (equivalent to RME $< \pm 5\%$) with high confidence (7).

Final Remark. We wholeheartedly agree with Kimbrough and Wakakuwa's assertion that "As the U.S. EPA is reconsidering its use of MDL, it should debate the importance of accuracy as well as precision in determining reporting limits". This is one of the major points of discussion in the current deliberations in ASTM D19 (water) subcommittee on detection. Departing from Kimbrough and Wakakuwa, we think traditional definitions of detection outcomes can and should still be used to demonstrate the inadequacies of the current 40CFR136b procedure. Fortunately, Hubaux and Vos have provided an improved MDL with detection performance that can be selected and an excellent starting point for a useful and statistically sound PQL.

Acknowledgments

The opinions expressed are those of the author and not necessarily those of Alcoa.

Literature Cited

(1) Kimbrough, D. E.; Wakakuwa, J. *Environ. Sci. Technbol.* **1993**, *27* (13), 2692–2699.
(2) EPA. *Appendix B to Part 136—Definition and procedure for the determination of the method detection limit, revision 1.11 (40CFR136)*; U.S. EPA: Washington, DC, 1990.
(3) Coleman, D. E. Some Simulation Results of Different Types of Method Detection Limits (MDLs). Publically available technical memo; Alcoa Labs, May 20, 1993.
(4) Hubaux, A.; Vos, G. Decision and detection limits for linear calibration curves. *Anal. Chem.* **1970**, *42* (8), 849–855.
(5) Liteanu, C.; Rica, I. *Statistical Theory and Methodology of Trace Analysis*; Ellis Horwood, Ltd.: Chichester, U.K., 1980; pp 96–104.
(6) EPA, procedure step 4 (a).
(7) Coleman, D. E. Technical Links Between Calibration, Relative Prediction Error (RPE), Detection Limits, Signal-to-noise, and ROCs. Publically available technical memo; Alcoa Labs, April 7, 1993; p 12.

David E. Coleman

Applied Mathematics & Computer Technology-D
100 Technical Drive
Alcoa Center, Pennsylvania 15069

ES940016N