Title
A Two-Component Model for Measurement Error in Analytical Chemistry

Permalink
https://escholarship.org/uc/item/1nz5x8rv

Journal
Technometrics, 37

Authors
Rocke, David
Lorenzato, Stefan

Publication Date
1995

TECHNOMETRICS, MAY 1995, VOL. 37, NO. 2

A Two-Component Model for Measurement Error in Analytical Chemistry

David M. Rocke
Graduate School of Management
University of California
Davis, CA 95616

Stefan Lorenzato
Division of Water Quality
California State Water Resources
Control Board
Sacramento, CA 94244-2130

In this article, we propose and test a new model for measurement error in analytical chemistry. Often, the standard deviation of analytical errors is assumed to increase proportionally to the concentration of the analyte, a model that cannot be used for very low concentrations. For near-zero amounts, the standard deviation is often assumed constant, which does not apply to larger quantities. Neither model applies across the full range of concentrations of an analyte. By positing two error components, one additive and one multiplicative, we obtain a model that exhibits sensible behavior at both low and high concentration levels. We use maximum likelihood estimation and apply the technique to toluene by gas-chromatography/mass-spectrometry and cadmium by atomic absorption spectroscopy.

KEY WORDS: Atomic absorption spectroscopy (AAS); Coefficient of variation; Detection limit; Gas-chromatography/mass-spectrometry (GC/MS); Maximum likelihood; Quantitation level.

INTRODUCTION: MEASUREMENT NEAR THE DETECTION LIMIT

Traditionally, the description of the precision of an analytical method is accomplished by applying two separate models, one for describing zero and near-zero concentrations of the analyte (the compound that the analytical method is designed to measure) and another for quantifiable amounts. This traditional approach leaves a "gray area," of analytical responses in which the precision of the measurements cannot be determined. The model governing quantifiable amounts assumes that the likely size of the analytical error is proportional to the concentration. If this model is applied to analytical responses in the gray area, then there is an implicit assumption that the analytical error becomes vanishingly small as the measurements approach 0. From long experience, this assumption appears to be invalid. Similarly, if the zero-quantity model is applied, there is an implicit assumption that the absolute size of the analytical error is unrelated to the amount of material being measured. Based on similar empirical information, this assumption also cannot be supported.

The new model presented in this article resolves these difficulties by providing an estimate of analytical precision that varies between the two extremes described by the traditional models. The model provides a distinct advantage over existing methods by describing the precision of measurements across the entire usable range. Examples are given of two different analytical methods, an atomic absorption spectroscopy analysis for cadmium and a gas-chromatography/mass-spectrometry analysis for toluene, both of which support the validity of the new model. The new model is applicable to a wide variety of situations including nonlinear calibration and added-standards calibration. Discussion is provided on the application of the new model to some common issues such as determination of detection limits, characterization of single samples, and determination of sample size required for inference to given tolerances.

Many measurement technologies have errors whose size is roughly proportional to the concentration—this is often true over wide ranges of concentration (Caulcutt and Boddy 1983). One common way to describe this constant coefficient of variation (CV) model is that the measured concentration x is given by

$$\log(x) = \log(\mu) + \eta$$

or

$$x = \mu e^{\eta},$$

where μ is the true concentration and η is a normally distributed analytical error with mean 0 and standard deviation σ_η. This model is widely used, but it fails to make sense for very low concentrations because it implies absolute errors of vanishingly small size.

On the other hand, one often considers the case of analyzing blanks (samples of zero concentration) and data near the detection limit by the model

$$x = \mu + \epsilon$$

with normally distributed analytical error ϵ. This is used for calibration with $\mu = 0$ and for the determination of detection limits (Massart, Vandingste, Deming, Michotte, and Kaufman 1988). It is, however, a bad approximation over wider ranges of concentration because it implies that the absolute size of the error does not increase with the concentration, and this is rarely true of analytical methods.

The solution proposed here is a combined model that reflects two types of errors. For example, in a gas-chromatography/mass-spectrometry (GC/MS) analysis one type of error is in the generation and measurement of peak area, which will generally have errors of size proportional to the concentration. The other type of error comes from the fact that, even when fed with a blank sample, the output is not a flat line but still retains some small variation. There are many sources of error in such an analysis; the idea here is merely to classify them into two types, additive and multiplicative. The model proposed is

$$x = \mu e^{\eta} + \epsilon, \tag{1.4}$$

where there are two analytical errors, $\eta \sim N(0, \sigma_\eta^2)$ and $\epsilon \sim N(0, \sigma_\epsilon^2)$, each normally distributed with mean 0. Here η represents the proportional error that is exhibited at (relatively) high concentrations and ϵ represents the additive error that is shown primarily at small concentrations. Another way of writing this model is via a linear calibration curve as

$$y = \alpha + \beta \mu e^{\eta} + \epsilon, \tag{1.5}$$

where y is the observed measurement (such as peak area a GC). Note that the new model approximates a constant standard deviation model for very low concentrations and approximates a constant CV model for high concentrations.

The approach proposed here should be contrasted with an alternative method, which is to model the standard deviation as a linear function of the mean concentration. The latter approach should work well over restricted ranges but suffers from some disadvantages when used over a wide range. First, the predicted standard deviation at zero concentration need bear little resemblance to the measured value when one regresses the standard deviation of replicates on the mean of the replicates. The proposed new model allows the data near 0 to determine the predicted standard deviation near 0 and the data for large concentrations to determine the standard error for large concentrations. Second, the new model allows the errors at large concentrations to be lognormal, rather than normal, which is in accord with much experience. Third, there is a more plausible physical mechanism for the existence of multiplicative and additive errors than there is for a standard deviation that is linear in the mean. A comparison of the predicted standard deviations from the two models is given in a later example.

If viewed in terms of the coefficient of variation, the picture is that large concentrations have a constant CV, whereas small concentrations have an increasing CV that tends to infinity as the concentration approaches 0. This bears some relationship to what Hall and Selinger (1989) called the "Horwitz trumpet" (Horwitz 1982; Horwitz, Kamps, and Boyer 1980). There are substantial differences however, that are discussed in Appendix A.

As an example, suppose that $\sigma_\epsilon = 1$ part per billion (ppb) and $\sigma_\eta = .1$. Then the standard deviation of blanks is 1 ppb, so the detection limit (in one definition) might be set at $3\sigma_\epsilon = 3$ ppb. At this concentration, measurements have a standard deviation of 1.04 ppb (using the new model), barely above the value at zero concentration—this is, a coefficient of variation of .35. One definition of the quantitation level is concentration in which the coefficient of variation falls to .2, the *practical quantitation level* (Environmental Protection Agency 1989). Because $\sigma_\eta = .1$, the coefficient of variation is .1 for large concentrations. Using the new model, the critical level at which the CV is .2 can be found by solving $(.2x)^2 = (.1x)^2 + (1)^2$ for x, which yields $x = 5.77$. Thus, at 6 ppb, the Environmental Protection Agency's practical quantitation level (PQL) has been reached. This calculation is only possible because of the new model—neither of the standard models can be used to compute the PQL.

It has been stated that values below the PQL do not yield useful quantitative information about the concentration of the analyte (e.g., Massart et al. 1988). The incorrectness of this idea has been pointed out (e.g., ASTM-D4210 1990 [ASTM 1987]). The new model more clearly demonstrates the usefulness of measurements below quantitation levels. The proposed model allows determination of the error structure of an analytical method near the detection limit and so provides an easy way to give precisions for such measurements, as well as the average of several such.

2. ESTIMATION

The approach we used is based on maximum likelihood estimation. An observed value x differs from the theoretical value μ because of the two errors η and ϵ, which are not directly observed. Any combination of η and ϵ that satisfies $\epsilon = x - \mu e^{\eta}$ is possible. Consequently, the likelihood associated with a set $(\sigma_\eta, \sigma_\epsilon)$ of parameters given a set of n measurements x_i with known concentrations μ_i is

$$\prod_{i=1}^{n} \int_{-\infty}^{\infty} \frac{1}{2\pi \sigma_\epsilon \sigma_\eta} e^{-\eta^2/(2\sigma_\eta^2)} e^{-(x-\mu e^{\eta})^2/(2\sigma_\epsilon^2)} \, d\eta. \tag{2.1}$$

Maximizing this likelihood leads to estimates of the necessary parameters σ_η, and σ_ϵ. More complex models, such as the estimation of a calibration curve, can be estimated in the same fashion, using maximum likelihood. For example, the calibration model (1.5) has likelihood

$$\prod_{i=1}^{n} \int_{-\infty}^{\infty} \frac{}{2\pi \sigma_\epsilon \sigma_\eta} e^{-\eta^2/(2\sigma_\eta^2)} e^{-(y-\alpha-\beta\mu e^{\eta})^2/(2\sigma_\epsilon^2)} \, d\eta. \tag{2.2}$$

Once estimates $\widehat{\sigma}_n$ and $\widehat{\sigma}_\epsilon$ have been derived, the precision of any measured value in the form (1.4) is (using the formula for the variance of a lognormal random variable)

$$\sqrt{\sigma_\epsilon^2 + \mu^2 e^{\sigma_\eta^2}\left(e^{\sigma_\eta^2} - 1\right)}, \tag{2.3}$$

which can be estimated by substituting estimated values for the parameters. This formula illustrates the key feature of the new model—when μ is small, the error is nearly constant in size; when μ is large, the size of the error is roughly proportional to μ. In the calibration form (1.5), the estimated concentration is

$$\widehat{\mu} = \widehat{\beta}^{-1}(y - \widehat{\alpha}) = \widehat{\beta}^{-1}(\alpha - \widehat{\alpha} + \beta\mu e^{\eta} + \epsilon). \tag{2.4}$$

If the covariance matrix $(\widehat{\alpha}, \widehat{\beta})$ is V, then a straightforward delta-method calculation shows that the variance of $\widehat{\mu}$ is

$$g'Vg + \sigma_\epsilon^2/\beta + \mu^2 e^{\sigma_\eta^2}\left(e^{\sigma_\eta^2} - 1\right), \tag{2.5}$$

where $g = (\beta^{-1}, -\mu\beta^{-1})$. Note that, when sample sizes are sufficiently large that α and β can be regarded as known, the $g'Vg$ term disappears, and this basically coincides with (2.3).

The computational approach used is outlined in Appendix B. From the usual maximum likelihood estimation theory, the estimates of the parameters are normal with asymptotic variances given by the negative inverse of the information matrix, which we use for the matrix V mentioned previously when needed.

2.1 Elaborations of the Basic Model

The model we are developing will be of no practical use unless it accurately describes the behavior of a wide range of analytical data. If a test data set consists of replicated measurements at a variety of levels from 0 to (say) 500 ppb, then the precision at 0 will be σ_ϵ and the CV at 500 ppb will be approximately σ_η, so these two parameters can be chosen to fit almost any behavior at the extremes. The model, if correct, says much more. From these two parameters, the exact way in which the standard deviation gradually rises in the transition region can be completely predicted. Thus it would make sense to compare the precision at each level implied by the model and the parameter fits to the estimated precision from the replicates. If the model in its simplest form does not fit the data, then an elaborated form may be required. One example of the possible need for an elaboration is a nonlinear calibration function. With a nonlinear calibration function $w(\mu;\theta)$, the model becomes

$$y = \alpha + w(\mu;\theta)e^{\eta} + \epsilon \tag{2.6}$$

The parameters can be determined by maximum likelihood as before.

Another problem that may arise is that the variance may not depend on the mean for large concentrations in the way implied by the model. An elaboration to help resolve this problem is to specify that $\eta \sim N(0, \sigma_\eta^2 V(\mu))$ (Carroll and Ruppert 1988). This allows for a different behavior than implied by lognormal errors.

A third potential elaboration is needed to address the case in which a labeled standard is used to calibrate measured values, which adjusts for recovery efficiency. Consider an analytical method to determine the concentration of a volatile organic such as toluene. One method of increasing accuracy is to spike the sample with a known concentration ν of deuterated toluene and determine the estimated concentration $\widehat{\mu}$ of toluene and $\widehat{\mu}_d$ of deuterated toluene. One estimate of the true concentration of toluene is $\widehat{\mu}_{\text{adj}} = \widehat{\mu}\nu/\widehat{\mu}_d$, which adjusts for recovery. If

$$y = \alpha + \beta\mu e^{\eta_1} + \epsilon \tag{2.7}$$

and

$$y_d = \alpha + \beta\nu e^{\eta_2} + \epsilon_2, \tag{2.8}$$

then the precision of $\widehat{\mu}_{\text{adj}}$ depends on the precision of $\widehat{\mu}$, the precision of $\widehat{\mu}_d$, and their covariance. Using the delta method we derive an approximate variance for $\widehat{\mu}_{\text{adj}}$ as

$$\operatorname{var}(\widehat{\mu}_{\text{adj}}) \approx \sigma_{\hat{\mu}}^2\nu^2/\widehat{\mu}_d^2 + \sigma_{\hat{\mu}}^2\nu^2\widehat{\mu}^2/\widehat{\mu}_d^4 - 2\sigma_{\hat{\mu}\hat{\mu}_d}\nu^2\widehat{\mu}/\widehat{\mu}_d^3, \tag{2.9}$$

where $\sigma_{\hat{\mu}}^2$, $\sigma_{\hat{\mu}_d}^2$, and $\sigma_{\hat{\mu}\hat{\mu}_d}$ are derived from a multivariate version of (2.5). In this case, it is necessary to have sufficient data to estimate the covariances of the errors in (2.7) and (2.8).

3. APPLICATIONS

In this section we describe some ways that the new model can be used. We concentrate especially on applications in environmental monitoring, where detection and measurement of low levels of toxic substances may be quite important. Generally, these applications assume that the parameters of the model have been determined, so we will generally describe the ideas in terms of the simple model (1.4). Extensions to the case in which calibration error is also to be accounted for are straightforward.

3.1 Detection Limits

According to the preceding model, the observations at true concentration $\mu = 0$ are normally distributed with standard deviation σ_ϵ. If r replicates are used, then any average of measured values greater than $D = 3\sigma_\epsilon/\sqrt{r}$ is extremely unlikely to have come from a zero concentration sample. Use of the exact method of setting confidence intervals described in Section 3.2 allows the precise determination of the uncertainty. Of course, this assumes that the replicates are true reruns of the entire process; otherwise the error may not be reduced by a factor of $\sqrt{r}$ but by a much smaller amount.

An implication of this rule for environmental monitoring is that the accumulation of measurements at low levels, even individually below the individual observation detection limit can still provide quantitative evidence of the concentration of a toxic substance. If the safe level is near or below the detection limit, then it might make sense to require replicate measurements (to reduce the effective detection limit) and to require the quantitative recording of measurements even when they are below the individual observation detection limit.

3.2 Uncertainty of a Single Measurement

There are two primary approaches to this problem, an exact solution and a normal or lognormal approximation. In the exact solution, we construct a confidence interval for the true concentration by locating the points on either side of the measured value where a hypothesis test just rejects. If a measurement x is obtained, and a 95% confidence interval is desired, then we need to find values μ_L and μ_U such that

$$\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma_\eta} \cdot e^{-\eta^2/(2\sigma_\eta^2)} \left(1-\Phi\left(x-\mu_L e^{\eta}\right)/\sigma_\epsilon\right) d\eta = .025 \tag{3.1}$$

and

$$\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma_\eta} e^{-\eta^2/(2\sigma_\eta^2)} \left(\Phi\left(x-\mu_U e^{\eta}\right)/\sigma_\epsilon\right) d\eta = .025, \tag{3.2}$$

where $\Phi()$ is the standard normal distribution function. This can be done by numerical solution of these equations, which yields a 95% confidence interval (μ_L, μ_U).

An approximate method is based on the estimated variance of x given by

$$V(x) = \sigma_\epsilon^2 + x^2 e^{\sigma_\eta^2}\left(e^{\sigma_\eta^2} - 1\right). \tag{3.3}$$

For low levels of x (those in which the first term dominates), the distribution of x is approximately normal. An approximate 95% confidence interval is formed as

$$x \pm 1.96\sqrt{V(x)}. \tag{3.4}$$

For high levels of x [those in which the second term in $V(x)$ dominates] $\ln(x)$ is approximately normally distributed with variance σ_η^2 so that a 95% confidence interval for μ is

$$(\exp(\ln(x) - 1.96\sigma_\eta), \exp(\ln(x) + 1.96\sigma_\eta)). \tag{3.5}$$

Note that this interval, although symmetric on the logarithmic scale, is asymmetric on the original measurement scale.

3.3 Uncertainty of an Average of Several Measurements

Here, the exact method of confidence-interval determination would be burdensome because it would require the computation of a difficult convolution, so confidence intervals must be based on the approximate normality of $\overline{x}$ (for low levels) or of $\overline{\ln(x)}$ (for high levels). For low levels, the average, $\overline{x}$, of n measurements will be approximately normally distributed with standard deviation $\sqrt{V(x)/n}$. For larger values of x, it will be better to perform inference using the logarithms of the data, which will be more nearly normally distributed. The average of the logarithms of n measurements will have approximate standard deviation $\sigma_\eta/\sqrt{n}$. Some further work will be needed to determine when inference should be based on the measurements and when on the logarithms. Another possible technique to improve inference is to use a normalizing transformation that depends on the parameter values.

Table 1. Cadmium Concentrations by AAS

Cadmium concentration (ppb)	Absorption (× 100)		
0	.0	−.7	−.1
2.7784	5.5	5.9	6.1
9.675	21.8	22.5	23.2
22.9716	53.4	53.6	50.9
31.7741	74.1	74.0	71.2
43.2067	94.6	99.6	99.4

3.4 Determination of Sample Size

The approximate sample size required to determine a concentration to a particular precision depends on the concentration as well as the precision desired. Suppose that we wish to choose the number of replicates r so that the chance of detecting a specified concentration that is above the safe level is sufficiently high. For example, suppose that the safe level is .1 ppb and the standard deviation parameters are $\sigma_\epsilon = .2$ ppb and $\sigma_\eta = .1$. Suppose that it was held to be important to detect a concentration of .3 ppb. At that concentration, the standard deviation of an average of r replicates is

$$\sqrt{\left[\sigma_\epsilon^2 + \mu^2 e^{\sigma_\eta^2}\left(e^{\sigma_\eta^2} - 1\right)\right]/r}$$

$$= \sqrt{[.04 + (.09)e^{.01}(e^{.01} - 1)]/r} \tag{3.6}$$

$$= \sqrt{.0409/r} \tag{3.7}$$

$$= .202/\sqrt{r}. \tag{3.8}$$

Using a normal approximation, the distance between the safe level .1 ppb and the concentration .3 ppb in standard deviation units is

$$(.2/.202)\sqrt{r} \quad .989\sqrt{r}.$$

Table 2. Standard Deviation of Absorption and Log-Absorption

Cadmium concentration	Absorption std. dev.	Log-absorptio std. dev.
0	.351	
2.7784	.283	.0488
9.675	.645	.0287
22.9716	1.360	.0260
31.7741	1.564	.0215
43.2067	2.821	.0289

Table 3. Initial Estimates From Pooled Data and Final Parameters by MLE for Cadmium Data

Parameter	Starting value	Final value
α	−.0963	−.3691
β	2.292	2.315
σ_η	.02511	.02507
σ_ϵ	.3188	.2970

For the chance of detection to be .95, we require $.989\sqrt{r} > 1.645$ or $r > 2.77$. In this case, the number of replicates should be at least 3.

4. EXAMPLES

In this section, we present two examples of analytical methods and examine the fit of the two-component model. The first example consists of graphite-furnace atomic absorption spectroscopy (AAS) measurements of cadmium, for known concentrations from 0 to 43 ppb, each quadruply replicated. The second example is GC/MS measurements of toluene from 4.6 picograms to 15 nanograms. For descriptions of these methods see Willard, Merritt, Dean, and Settle (1988). In both cases we use the calibration form (1.5) with likelihood (2.2).

4.1 Cadmium by Atomic Absorption Spectroscopy

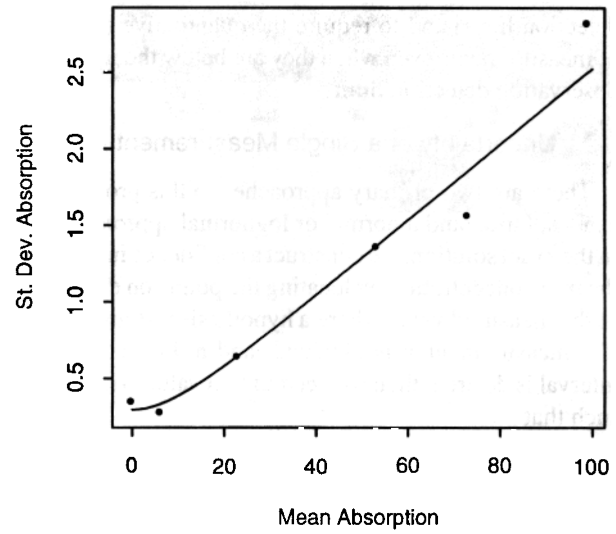

Figure 2. Actual and Predicted Standard Deviation for Cadmium by AAS. The dots represent the standard deviation of the replicates and the solid line is the predicted standard deviation using (3.4) and (3.5) as appropriate.

The instrument used for these measurements was a Perkin-Elmer model 5500 graphite furnace atomic absorption spectrometer with a model HGA-500 furnace controller. The tubes were pyrolytically coated; injections were made onto L'vov platforms. A deuterium arc lamp was used for background correction. Table 1 shows the AAS measurements, and Table 2 gives the standard deviation of the replicate measurements of absorbance and the standard deviation of the logarithms of the replicates. Note that the measurements at the two lowest concentrations appear to have a constant standard deviation of the measured concentration and the measurements at the four highest concentrations appear to have a constant standard deviation of the measured log-concentration. Therefore, the data span the "gray area" of quantification. Starting values for the maximum likelihood estimator (MLE) procedure, which in this case are themselves quite good

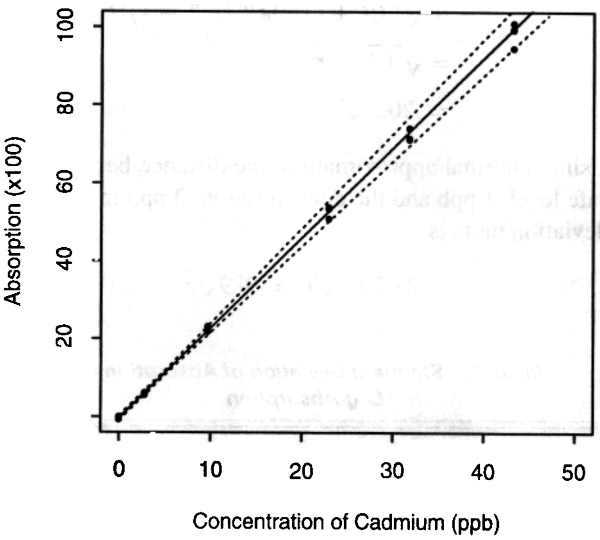

Figure 1. Concentration of Cadmium Versus Absorption by AAS. The dots represent the measured values, the solid line is the calibration curve estimated by maximum likelihood using model (1.5), and the dashed lines form an estimated prediction envelope using (3.4) and (3.5) as appropriate.

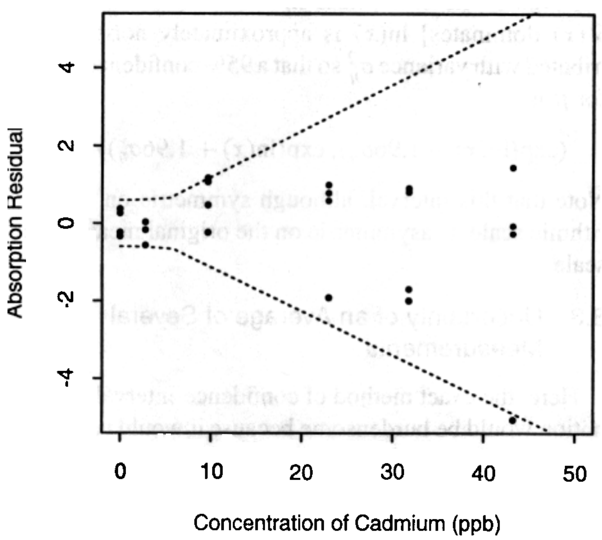

Figure 3. Residuals for Cadmium Calibration. The dots represent the differences between measured values and the predicted values from Model (1.5) and the dashed lines form an estimated prediction envelope using (3.4) and (3.5) as appropriate.

Table 4. Toluene Amounts by GC/MS

Toluene amount (pg)	Peak area			
4.6	29.80	16.85	16.68	19.52
23	44.60	48.13	42.27	34.78
116	207.70	222.40	172.88	207.51
580	894.67	821.30	773.40	936.93
3,000	5350.65	4942.63	4315.79	3879.28
15,000	20718.14	24781.61	22405.76	24863.91

Table 6. Initial Estimates From Pooled Data and Final Parameters by MLE for Toluene Data

Parameter	Starting value	Final value
α	1.6	11.51
β	1.546	1.524
σ_η	.10	.1032
σ_ϵ	6.0	5.698

estimates of the parameters, are shown in Table 3. Starting values for the regression parameters α and β were derived by ordinary linear regression of the absorbance measurements on the concentration. The starting value for the standard deviation σ_ϵ of zero measurements was estimated from the two lowest concentrations because there appears not to be an upward trend until the third highest concentration. A starting value for the standard deviation of the log absorbances for high concentrations σ_η, which is approximately the coefficient of variation, was derived by pooling the highest four levels, where the CV appears constant. These initial estimates are of such high quality that they are scarcely changed by the MLE iterations. Final values are also shown in Table 3. Note that the calculation is not extremely sensitive to the starting value because the same optimum was arrived at from $\alpha = 0, \beta = 2, \sigma_\eta = .03$, and $\sigma_\epsilon = .4$, for example. Somewhat plausible values must be used, however, to avoid numerical instability.

The results are shown graphically in Figure 1, which plots the data, the calibration curve, and an estimated envelope; Figure 2, which shows the actual and predicted standard deviation; and Figure 3, which shows the residuals from the model fit.

If we now treat these estimated parameters as known, we can use the model for further analysis. For example, if the limit of detection is defined as the absorbance (and associated implied concentration) that falls three standard deviations above 0, we find that it lies at about .4 ppb.

Confidence intervals for concentration can be derived using the methods of Section 3.2. For example, with the parameters estimated for the cadmium data, an absorbance of 6 implies an estimated concentration of 2.75 ppb, with 95% confidence interval (2.47, 3.04). This is almost symmetric, with half widths of .28 and .29. An absorbance of 50 implies an estimated concentration of 21.76 ppb, with 95% confidence interval (20.69, 22.88) and half-widths 1.07 and 1.12. A normal approximation using Equation (2.3) yields confidence intervals of (2.49, 3.01) for an absorbance of 6 and (21.23, 22.29) for an absorbance of 50. The approximation is quite acceptable for the low absorbance but not for the higher one. For the absorbance of 50, an acceptable approximation is gained by using the approximate lognormality of the measurements at high levels, leading to a 95% confidence interval of (20.72, 22.85). Note, however, that there is no satisfactory alternative to the exact confidence interval provided by the new model that is accurate over the entire range of measurements.

Table 5. Standard Deviation of Peak Area and Log Peak Area

Toluene amount	Peak area std. dev.	Log peak area std. dev.
4.6	6.20	.2719
23.0	5.65	.1387
116.0	21.02	.1080
580.0	73.19	.0858
3,000.0	652.98	.1427
15,000.0	2005.02	.0878

4.2 Toluene by GC/MS

The instrument used here was a Trio-2 GC/MS (VG Masslab) with electron ionization at 70 electron volts with a 30-meter DB-5 GC column. Table 4 shows an analysis of amount of toluene by GC/MS for known amounts of from 4.6 picograms to 15 nanograms in 100 μL of extract. (1 picogram in 100 μL corresponds to a concentration of .01 ppb.) The quantitation is done by peak area

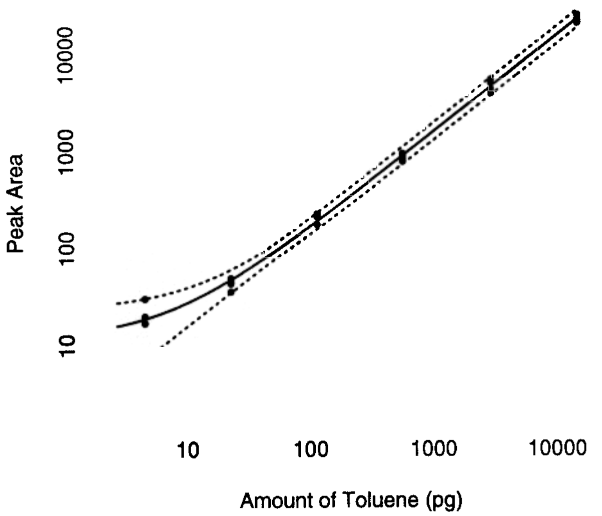

Figure 4. Concentration of Toluene Versus Peak Area by GC/MS. The dots represent the measured values, the solid line is the calibration curve estimated by maximum likelihood using Model (1.5), and the dashed lines form an estimated prediction envelope using (3.4) and (3.5) as appropriate. Note the log-log scale.

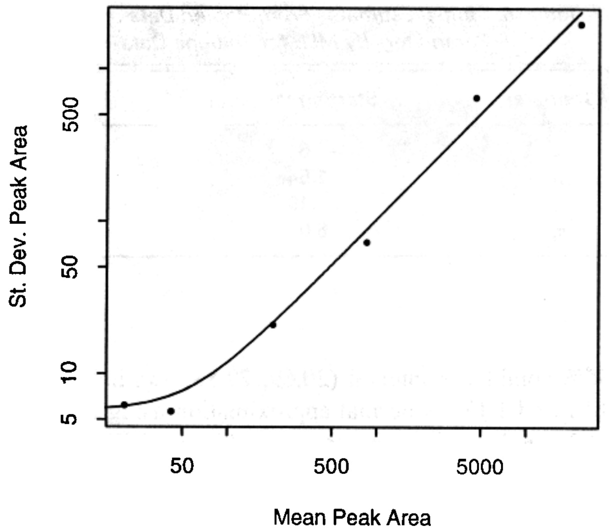

Figure 5. Actual and Predicted Standard Deviation for Toluene by GC/MS. The dots represent the standard deviation of the replicates and the solid line is the predicted standard deviation using (3.4) and (3.5) as appropriate. Note the log-log scale.

at m/z 91. The relationship between amount and peak area is satisfactorily linear and the behavior of the errors is generally consistent with the new model as shown in Table 5. Starting values were derived from ordinary linear regression and from an examination of the standard deviations of the raw and logged data. The MLE estimation gives apparently acceptable results as shown in Table 6.

The results are shown graphically in Figure 4, which plots the data, the calibration curve, and an estimated envelope; Figure 5, which shows the actual and predicted standard deviation; and Figure 6, which shows the residuals from the model fit. These plots are all on a log-log scale.

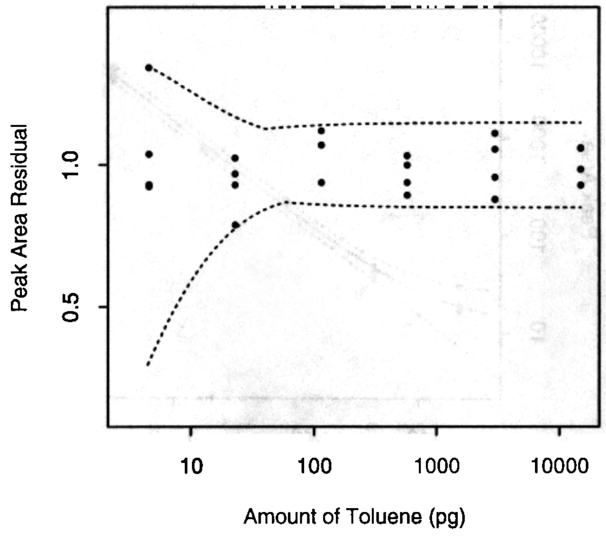

Figure 6. Residuals for Cadmium Calibration. The dots represent the ratios between measured values and the predicted values from Model (1.5) and the dashed lines form an estimated prediction envelope using (3.4) and (3.5) as appropriate. Note the log-log scale.

Table 7. Predicted Standard Deviations From New Model and From Linear Regression of the Standard Deviation on the Mean

Mean peak area	*Peak area std. dev.*	*Predicted std. dev. from (2.5)*	*Predicted std. dev. from linear regression*
20.7	6.20	5.74	46.60
42.4	5.65	6.76	48.48
202.6	21.02	19.25	62.29
856.6	73.19	92.13	118.68
4622.1	652.98	475.65	443.37
23192.4	2005.02	2378.08	2044.64

A comparison is given in Table 7 between the predicted standard deviation of the peak area from the model developed in this article and the predicted standard deviation from a variance function in which the standard deviation is a linear function of the mean. Both fit well for high concentrations, but the linear function is quite inaccurate for lower concentrations. For these large ranges, it appears that the new model fits the behavior of the data better.

5. CONCLUSION

In this article we have provided a new model for analytical error that behaves like a constant standard deviation model at low concentrations and like a constant CV model at high concentrations. The importance of this new model is that it provides a reliable way to estimate the precision of measurements that are near the detection limit so that they can be used in inference and regulation. Although the illustrations used in the article are for linear calibration analyses, the model itself is flexible enough to be used in a wide variety of situations, including nonlinear calibration, censored data, and added-standards methods.

ACKNOWLEDGMENTS

We are grateful to A. Russell Flegal, Institute of Marine Sciences, University of California, Santa Cruz, and A. Daniel Jones, Facility for Advanced Instrumentation, University of California, Davis, for providing the data used as examples in this article, and to the editor, an associate editor, and two referees for suggestions that substantially improved the article. The research reported in this article was supported in part by the California State Water Resources Control Board, by National Institute of Environmental Health Sciences, National Institutes of Health Grant P42 ES04699, and by National Science Foundation Grant DMS-93.01344.

APPENDIX A: THE HORWITZ TRUMPET

Horwitz (1982; Horwitz et al. 1980) examined over 150 independent Association of Official Analytical Chemists interlaboratory collaborative studies covering numerous topics from drug preparations and pesticide formulations (at the high concentration end) to aflatoxin (at the low end). Over these studies, there was a pattern to the interlaboratory CV that Hall and Selinger (1989) described as the Horwitz trumpet. The average CV of the interlaboratory errors increased from a few percent for concentrations in the double-digit percentage range, to about 50% for concentrations in the 1 ppb range. Horwitz suggested an empirical formula relating the concentration μ to the coefficient of variation $\mathrm{CV}(\mu)$ [and therefore the precision $\sigma(\mu)$] as follows:

$$\mathrm{CV}(\mu) = .006\mu^{-.5} \qquad \text{(A.1)}$$

and

$$\sigma(\mu) = .006\mu^{.5}. \qquad \text{(A.2)}$$

After an analysis of Horwitz's data using a binomial model with variable apparent sample size, Hall and Selinger suggested a similar formula:

$$\mathrm{CV}(\mu) = .02\mu^{-.15} \qquad \text{(A.3)}$$

and

$$\sigma(\mu) = .02\mu^{.85}. \qquad \text{(A.4)}$$

Both the Horwitz trumpet and the present article suggest an increase in the CV at low concentrations; however, there are several important differences. First, Horwitz was interested in the variation from laboratory to laboratory, whereas the model in this article is intended for intralaboratory precision (there is no reason why it should not apply also across laboratories, however). Second, and more important, Horwitz was describing how the precision changes when one moves from one analytical method intended for a certain range of concentrations to another method intended for a different range. The two-component error model is for a single analytical method across its useful range. There is no reason why one model cannot be used within methods and a different one between methods. In the language of this article, Horwitz had a model for how σ_η changes with the analytical method, but the two-component model describes how the precision changes within a given method as the concentration approaches the detection limit of that method.

It is of particular note that the Horwitz trumpet cannot be used to serve the purpose of this article in describing the transition of the error structure in a given analytical method as the concentration changes from high levels to those near the detection limit. This is because the Horwitz model in both its original form and in Hall and Selinger's emendation imply that the standard deviation of a measurement at low levels approaches 0. In particular, if used inappropriately to describe the interlaboratory error under zero concentration, this model would imply zero error. In fact, most laboratories might report the compound as "not detected," but this is a far cry from zero error.

In summary, both the Horwitz trumpet and the two-component error model have a proper place in understanding the errors of analytical method. The former is useful to describe the way in which the size of the errors from analytical methods change as the method itself and the intended range of concentration changes. The latter, the focus of this article, allows better estimation of quantitation error for a given analytical method for measurements within a few multiples of the detection limit.

APPENDIX B: COMPUTATIONAL ISSUES IN ESTIMATION

Actually performing the maximization of the likelihood requires the numerical optimization of a numerically integrated likelihood. For the numerical integration, we used a Gauss–Hermite rule with centering depending on a second-order Taylor approximation to the logarithm of the integrand. This led to highly accurate results. Specifically, we need to approximate

$$\int_{-\infty}^{\infty} \frac{1}{2\pi\sigma_\epsilon\sigma_\eta} e^{-\eta^2/(2\sigma_\eta^2)} e^{-(x-\mu e^\eta)^2/(2\sigma_\epsilon^2)}\,d\eta. \qquad \text{(A.5)}$$

Gauss–Hermite integration is a form of Gaussian integration for use with a kernel of the form $w(x) = \exp(-x^2/2)$ (Davis and Rabinowitz 1984). One approximates an integral

$$\int f(x)w(x)\,dx \qquad \text{(A.6)}$$

by

$$\sum^{m} w_i f(x_i), \qquad \text{(A.7)}$$

where the Gaussian points x_i and weights w_i can be obtained from a variety of computer programs. (Davis and Rabinowitz 1984; IMSL Library 1989; NAG Library 1987). This is exact when $f(x)$ is a polynomial of degree $2m+1$ or less.

The Equation (A.5) can be transformed to resemble (A.6) by a linear transformation. In fact, we transform (A.5) so that it resembles $w(x) = \exp(-x^2/2)$ by matching the first two derivatives of the logarithm at the mode. A function like

$$g(x) = e^{-(x-x_0} \qquad \text{(A.8)}$$

satisfies

$$\frac{d\log(g(x))}{dx} = (x \quad x_0)/k \qquad \text{(A.9)}$$

and

$$\frac{d^2\log(g(x))}{dx^2} = -1/k^2. \qquad \text{(A.10)}$$

Then x_0 is a 0 of $d\log(g(x))/dx$, and

$$k = -\left(\frac{d^2\log(g(x))}{dx^2}\right)^{-1/2}_{x=x_0} \qquad \text{A}$$

Thus, if we match the integrand of (A.5) in this fashion and transform using $\eta^* = (\eta - \eta_0)/k$, we approximate the form of the kernel.

Now the log of the integrand of (A.5) is

$$h(\eta) = -\log(2\pi\sigma_\epsilon\sigma_\eta) - \eta^2/(2\sigma_\eta^2) - (x - \mu e^\eta)^2/(2\sigma_\epsilon^2). \tag{A.12}$$

$$h'(\eta) = -\eta/\sigma_\eta^2 + \mu e^\eta(x - \mu e^\eta)/\sigma_\epsilon^2 \tag{A.13}$$

and

$$h''(\eta) = -1/\sigma_\eta^2 + \mu e^\eta(x - \mu e^\eta)/\sigma_\epsilon^2 - \mu^2 e^{2\eta}/\sigma_\epsilon^2. \tag{A.14}$$

We find a root of Equation (A.13) numerically and use the associated transformation to approximate the integral.

This procedure results in fairly quick calculation of an accurate approximation of the likelihood for use in maximum likelihood estimation. In the calculations in this article, we use a 12-point approximation. The numerical optimization was performed with a quasi-Newton method using BFGS rank-two updates and a trust region approach (Dennis and Schnabel 1983). The implementation used was the IMSL routine DUMINF (1989). Starting values for the iterations when the model is in the simple form (1.4) can be derived from two simple calculations. First, σ_ϵ can be estimated by the standard deviation of zero or near-zero replicates. Then, σ_η can be estimated by the standard deviation of the logarithms of some high concentration. For the calibration form (1.5), initial estimates of α and β can be derived from linear regression of the response on the log concentration.

[Received June 1992. Revised July 1994.]

REFERENCES

ASTM (1987), "Standard 4210," in *American Society for Testing and Materials 1987 Book of Standards*, Philadelphia: Author.

Carroll, R. J., and Ruppert, D. (1988), *Transformation and Weighting in Regression*, New York: Chapman and Hall.

Caulcutt, R., and Boddy, R. (1983), *Statistics for Analytical Chemists*, London: Chapman and Hall.

Davis, P. J., and Rabinowitz, P. (1984), *Methods of Numerical Integration* (2nd ed.), Orlando, FL: Academic Press.

Dennis, J. E., and Schnabel, R. B. (1983), *Numerical Methods for Unconstrained Optimization and Nonlinear Equations*, Englewood Cliffs, NJ: Prentice-Hall.

Environmental Protection Agency (1989), "Proposed Rule: Method Detection Limits and Practical Quantitation Levels," *Federal Register*, 54: 97, 22,100–22,104.

Hall, P., and Selinger, B. (1989), "A Statistical Justification Relating Interlaboratory Coefficients of Variation With Concentration Levels," *Analytical Chemistry*, 61, 1465–1466.

Horwitz, W. (1982), "Evaluation of Analytical Methods Used for Regulation of Food and Drugs," *Analytical Chemistry*, 54, 67A–76A.

Horwitz, W., Kamps, L. R., and Boyer, K. W. (1980), "Quality Assurance in the Analysis of Foods for Trace Constituents," *Journal of the Association of Official Analytical Chemists*, 63, 1344–1354.

IMSL Library (1989), Houston, TX: IMSL Inc.

Massart, D. L., Vandingste, B. G. M., Deming, S. N., Michotte, Y., and Kaufman, L. (1988), *Chemometrics: A Textbook*, Amsterdam: Elsevier.

NAG Library (1987), Oxford, U.K.: Numerical Algorithms Group, Ltd.

Willard, H. H., Merritt, L. L., Jr., Dean, J. A., and Settle, F. A., Jr. (1988), *Instrumental Methods of Analysis* (7th ed.), Belmont, CA: Wadsworth.