

Application to Environmental Monitoring of a Two-Component Model for Chemical Analytical Error

Machelle Wilson
University of California, Davis

David M. Rocke
University of California, Davis

Blythe Durbin
University of California, Davis

Henry D. Kahn
United States Environmental Protection Agency

June 2000

The use of analytical chemistry measurements in environmental monitoring is dependent on an assessment of measurement error. This paper demonstrates the utility of a two-component error model in setting limits of detection and related concepts. The model is applicable to analytical methods in which high concentrations are measured with approximately constant relative standard deviation. At low levels, the relative standard deviation cannot stay constant, since this implies vanishingly small absolute standard deviation. The two-component model has approximately constant standard deviation near zero concentration, and approximately constant relative standard deviation at high concentrations, a pattern that is frequently observed in practice.

1. Introduction

Limitations of the analytical methodology used to measure the concentration of toxic substances in the environment have had an important policy role in regulation. It is difficult to regulate emissions of toxic substances at levels below what can be reliably measured, and a definition of the level of reliable measurement is therefore crucial to policy making. In this paper, we discuss the implications of a model for measurement error for these policy issues.

It has been observed from long experience that the measurement error of an analytical method, for example atomic absorption spectroscopy, is of two types. Over a range of concentrations near zero, the measurement error is seen to be constant. Over ranges of higher concentration, the measurement error is observed to be proportional to the concentration of the analyte (Currie, 1968; Hubaux and Vox, 1970). This poses some difficulty in estimating the overall precision of an analytical method for data that span the “gray area” where a transition occurs between near-zero concentrations and quantifiable amounts. If a model assuming a constant error is used, there is an implicit assumption that the absolute size of the error is unrelated to the concentration of the analyte. This assumption is not supported by empirical observation of behavior at the higher levels. If a model assuming

a proportional error is used, then there is an implicit assumption that the measurement error becomes vanishingly small as the concentration approaches zero. This assumption is also contrary to experience of behavior at the lower levels. Since environmental monitoring data often fall into this gray area, understanding measurement error in this region is of considerable importance.

The model presented here resolves these difficulties by incorporating both types of error that are observed in practice into a single model. This model provides an obvious advantage over existing models by describing the precision of measurements across the entire usable range. We will present two examples—one of zinc by ICPMS and one of propionitrile by GC/MS. These examples support the validity and advantages of this two-component model. We also discuss the application of the model to detection limits, quantification limits, sample size calculations, and the construction of confidence intervals. Some of the technical background to this model can be found in Rocke and Lorenzato (1995). Some related work with a different point of view on some of the policy issues can be found in Zorn, Gibbons, and Sonzogni (1997, 1999).

2. The Model

Most measurement technologies require a linear calibration curve to estimate the actual concentration of an analyte in a sample for a given response. We can incorporate into the linear calibration model the two types of errors that are observed in most analyses. The two-component model is

$$y = \alpha + \beta\mu e^\eta + \epsilon \quad (2.1)$$

where y is the response (such as peak area) at concentration μ , $\eta \sim N(0, \sigma_\eta)$ and $\epsilon \sim N(0, \sigma_\epsilon)$. Here, η represents the proportional error that always exists, but is noticeable at concentrations significantly above zero, and ϵ represents the additive error that always exists but is noticeable mainly for near-zero concentrations. This two-component model approximates a constant standard deviation for very low concentrations and approximates a constant relative standard deviation (RSD) for higher concentrations. Note that y is the response of the measuring apparatus, for example peak area. In order to obtain the estimated concentration, we perform the back calculation:

$$\hat{\mu} = \frac{y - \hat{\alpha}}{\hat{\beta}}, \quad (2.2)$$

where $\hat{\alpha}$ and $\hat{\beta}$ are estimates of α and β respectively in the model (2.1).

Under this model, the variance of the response y at concentration μ is given by

$$\text{Var}\{y\} = \mu^2 \beta^2 e^{\sigma_\eta^2} (e^{\sigma_\eta^2} - 1) + \sigma_\epsilon^2 \quad (2.3)$$

(Rocke and Lorenzato 1995). Two derived quantities will be useful in interpretation of the results. $S_\epsilon = \sigma_\epsilon/\beta$ represents the approximate standard deviation of $\hat{\mu}$ at low levels

(approximate because it ignores uncertainty in the calibration parameters α and β). $S_\eta = \frac{\sigma_\eta}{\mu} e^{\sigma_\eta^2} (e^{\sigma_\eta^2} - 1)$ is the approximate RSD of $\hat{\mu}$ for high levels. For values of σ_η appropriate for analytical technologies (say no more than 0.3), S_η is very near σ_η . For example, if $\sigma_\eta = .1$, then $S_\eta = .1008$ and if $\sigma_\eta = .3$, then $S_\eta = .32$.

Using these derived quantities, we can represent the variance of y as

$$\text{Var}\{y\} = \mu^2 \beta^2 S_\eta^2 + \sigma_\epsilon^2 \quad (2.4)$$

and for the estimated concentration,

$$\text{Var}\{\hat{\mu}\} = \mu^2 S_\eta^2 + S_\epsilon^2 \quad (2.5)$$

Let us compare the usefulness of two-component error model to a model using only the relative standard deviation (RSD), which is defined to be equal to the standard deviation of the estimated concentration divided by the concentration (Liteanu and Rica, 1980). For the two component model, we have

$$\text{RSD}\{\hat{\mu}\} = \frac{\sigma_\epsilon}{\mu \sqrt{S_\eta^2 + S_\epsilon^2/\mu^2}} \quad (2.6)$$

If the error structure is described only in terms of RSD, we see that measurements at high concentrations have a nearly constant RSD, whereas small concentrations have an increasing RSD that tends to infinity as the concentration approaches zero. Use of RSD alone to characterize measurement error in the low concentration region can cause difficulties when attempting to make decisions regarding detection and quantification. The two-component model allows for a more reasonable estimation of error near zero, and hence more reasonable criteria for setting detection limits, critical levels, and quantification levels.

As an example (zinc by ICPMS described in detail later), suppose that $\sigma_\epsilon = 204$ (in units of the peak area), $\sigma_\eta = 0.0390$ (this is the high-level CV), $\alpha = 490$, and $\beta = 7.06$. The derived quantities are $S_\epsilon = 204/7.06 = 28.9\text{ppt}$ and $S_\eta = .0390$. Then the standard deviation of blanks is 204 (units of peak area) or 28.9ppt. Critical levels, which are the basis for determining detection of an analyte, are often set at 2–3 times the standard deviation of the blank above background (see Currie, 1995, 1997). Using this definition with a multiplier of 3, we have the critical level set at $490 + 3(204) = 1102$ in units of peak area or $3(28.9\text{ppt}) = 86.7\text{ppt}$. Then, using (2.4) the standard deviation of the response, y , at concentration $\mu = 86.7$ ppt is

$$\begin{aligned} \text{SD}\{y\} &= \sqrt{(86.7)^2(7.06)^2(.0390)^2 + (204)^2} \\ &= \sqrt{42,200} \\ &= 205 \end{aligned}$$

(compared to 204 at zero concentration). The standard deviation of the estimated concentration is, using (2.5),

$$\begin{aligned} \text{SD}\{\hat{\mu}\} &= \sqrt{(86.7)^2(.0390)^2 + (28.9)^2} \\ &= \sqrt{847} \\ &= 29.1\text{ppt} \end{aligned}$$

(compared to 28.9 ppt at zero concentration).

That is, using the two-component model, measurements at this concentration have a standard deviation of about 29.1 ppt, only slightly above the value for blanks, and $RSD = 29.1/86.7 = 0.34$.

3. Estimation

The parameters in the two-component model can be estimated in a number of ways. The standard deviation σ_ϵ of the low level measurements can be estimated from replicate blanks, for example by those routinely included with batches of samples. If this number is stable, it can also be estimated from routine QC data so long as measurements are replicated. The parameter σ_η can be likewise estimated from the standard deviation of the logarithm of high level measurements. The calibration curve can then be estimated using weighted least squares, weighting each point by the inverse estimated variance. It is also possible to estimate all four parameters simultaneously using weighted least squares, although our experience is that this estimation method is often not very stable and can lead to nonconvergence or impossible estimates (such as negative variances).

The most effective estimation method is maximum likelihood, as described in Rocke and Lorenzato (1995). A computer program that solves for the maximum likelihood estimates for α , β , σ_ϵ , σ_η is available at <http://handel.cipic.ucdavis.edu/~dmrocke>.

4. Applications

In this section we describe some applications of the two-component model. Special emphasis is given to applications in environmental monitoring, where detection and measurement of toxins at very low levels can be quite important.

4.1. Critical Levels

Detection refers to the capability of an analytical measurement process to indicate the presence of an analyte. This requires an agreed upon procedure for determining whether or not a given measurement result conclusively establishes that the analyte is present in the sample. In practice, this means that the investigators establish a numerical value such that a result greater than this value is extremely unlikely to occur if the true concentration is in fact zero, while a result lower than this value indicates that the true concentration in the sample is either zero, or is too low to detect (with certainty) with the technology in use. The measurement error that exists in any technology leads to this inability to detect concentrations below a certain level. The Critical Level is defined by the International Union of Pure and Applied Chemistry (IUPAC) to be the value, L_C , such that the probability of a measurement exceeding this value will be very small, say 0.01, when the true concentration in the sample is zero (Currie, 1995). That is, samples that do not contain the analyte are very unlikely to generate measurement results that exceed L_C . Note that the Critical Level

is defined at first in the units of the measurement technology (e.g., peak area) not in units of concentration. Of course, we can also express the critical level in units of concentration by taking the critical level (in measured units) and dividing by the calibration slope β . Since the Critical Level is the point at which the detection decision is made, it has been called by some authors a detection limit, but it should be noted that it is distinct from the IUPAC definition of the Limit of Detection. Limits of Detection will be discussed later in this paper.

Under the assumption of normality, the value of L_C may be calculated as follows: Assume that σ_ϵ , the standard deviation of the response at $\mu = 0$, is known and that we require 99% confidence in our statement that the analyte is present. Then the one-sided 99% confidence level is represented by $L_C = \alpha + z_0\sigma_\epsilon$, where z_0 is the z -value corresponding to the 99th percentile of the standard normal distribution (i.e., $z_0 = 2.326$). To find the critical level for any level of confidence, simply find the appropriate one-sided z -value, then multiply by the standard deviation of the blanks and add this to the mean value of the blanks i.e.,

$$L_C = \alpha + z_0\sigma_\epsilon \quad (4.7)$$

in units of the response of

$$L_C = z_0S_\epsilon \quad (4.8)$$

in units of concentration. Generally, the mean and standard deviation of the blanks will be well enough known from experience to use this method. If these are estimated from data, then a t -value (from the t -distribution), with the appropriate degrees of freedom, is substituted for the z -value. The advantage of the two-component error model is that an estimate for σ_ϵ with desirable statistical properties can be obtained from data that span a range of concentrations, resulting in greater accuracy from a given amount of data.

For our example zinc data, we have $\sigma_\epsilon = 204$ in units of peak area and $S_\epsilon = 28.9$ ppt. If we use a 99% confidence level, the normal percentage point is 2.326, so that L_C in units of peak area is $490 + (2.326)(204) = 965$ and in units of concentration is $(2.326)(28.9) = 67.2$ ppt.

Note that a measured value below L_C does not establish that the analyte is absent, only that it has not been shown conclusively to be present. This means that the value should be reported as measured, together with the standard deviation at the measurement value. Cases in which limitations of the instrument itself prevent reporting a value (e.g. as is the case with some spectroscopic measurements) are an obvious exception. In other cases, censoring of values below L_C may be required as a matter of policy by regulatory agencies. Such practices result in data sets with reduced utility from the loss of information, which makes tracking of trends, monitoring of laboratory quality, summarization of data, and other data analysis all more difficult. More importantly, this censoring needlessly prevents investigators from being able to reach probabilistically quantified conclusions about the true presence of an analyte. Sometimes sophisticated methods can be used to cope with these censored data (Gilbert 1987, Cohen 1991), but simple reporting of measured values would allow use of basic, easily understood methodology instead of these complex techniques.

4.2. Minimum Detectable Value

The Limit of Detection or Minimum Detectable Value is the true concentration, L_D , of an analyte that will, with high confidence, produce a measured value above the critical level. For example, if the concentration L_D is chosen for laboratory QC, it should be detected (measured above the Critical Level L_C) almost all of the time. Although L_C can be given either in the units of the measurement technology or in units of concentration, L_D is purely in units of concentration. Conceptually, L_C is determined so that the desired confidence level of the test that the true concentration is zero is met, and L_D is determined so that the desired statistical power is obtained. It usually cannot be safely assumed that the standard deviation at the detection limit is the same as the standard deviation of a blank, so reliable estimates of variance at any specified concentration are necessary for reliable determination of the minimum detectable value.

We can find a good estimate of L_D by noting that the level is low enough that a normal approximation is appropriate (at high levels, the distribution is essentially log normal). We treat y as being normally distributed with mean $\alpha + \beta\mu$ and with variance given by (2.4) and solve the resulting equation. Recall that, from (2.5),

$$\text{Var}\{\hat{\mu}\} = \mu^2 S_\eta^2 + S_\epsilon^2. \quad (4.9)$$

so that L_D is the solution to the equation

$$L_D = z_0 S_\epsilon^2 + z_1 \frac{\text{p} \overline{\text{Var}\{L_D\}}}{\text{q} \overline{\text{Var}\{L_D\}}} \quad (4.10)$$

$$= z_0 S_\epsilon^2 + z_1 \frac{L_D^2 S_\eta^2 + S_\epsilon^2}{L_D^2 S_\eta^2 + S_\epsilon^2}, \quad (4.11)$$

where z_0 is the percentile of the standard normal distribution corresponding to the desired confidence level for the critical level, and z_1 is the percentile corresponding to the desired confidence level for the minimum detectable value. Since L_D appears on both sides of the equation, at first glance it might seem that iterative methods would be required to find the solution. However, so long as the variance of y does not increase too rapidly with μ , a closed form solution is possible. The required condition is

$$S_\eta < 1/z_1. \quad (4.12)$$

For details and proof see the Appendix. The solution derived there is

$$L_D = \frac{S_\epsilon^2 z_0 + \frac{\text{h} \text{ q} \overline{\text{Var}\{L_D\}}}{z_0^2 - (1 - z_1^2 S_\eta^2)(z_0^2 - z_1^2)}}{(1 - z_1^2 S_\eta^2)}. \quad (4.13)$$

When $z_0 = z_1 = z$, we have the particularly simple form

$$L_D = \frac{2z S_\epsilon^2}{1 - z^2 S_\eta^2} \quad (4.14)$$

Then an estimated L_D is found by simply substituting the sample variance estimates into the above equation.

For our example zinc data, we have $S_\epsilon = 28.9$ and $S_\eta = .0390$. If we use $z_0 = z_1 = 2.326$, corresponding to 99% confidence, we have a minimum detectable value of

$$(2)(2.326)(28.9)/[1 - (2.326)^2(.0390)^2] = 135\text{ppt.} \quad (4.15)$$

One important use of L_D is to assess and monitor the performance of a laboratory. If samples are spiked at a concentration of L_D , then almost all of the time the resulting responses should exceed the critical value L_C . Such trials can be run periodically to monitor the ability of the laboratory to detect analytes up to specifications. Another important use is to determine what concentrations in the field can reliably be detected with a given technology. If concentrations below the minimum detectable value are important to detect, consideration should be given to the use of better technology or replicate measurements. The minimum detectable value should never be used to assess a measured value to decide if it should be reported or censored. It should only be used for planning purposes or for quality assurance.

4.3. Quantification Limit

By a Quantification Limit, some authors have meant the lowest level at which the quantitative assessment is sufficiently accurate for practical use. Since the standard deviation at low levels is actually smaller than that at high levels, it could be argued that this is a meaningless concept—any measured value, along with a confidence interval, provides a useful measurement; and the most precise measurements, in terms of standard deviation, are actually those for the lowest level of the analyte. A definition with some practical utility is the true concentration at which the relative standard deviation (RSD) falls to a specified level (Gibbons, 1994; Meier and Zünd, 1993). However, measurement of some analytes at an arbitrarily low RSD, such as 10% may not be possible. The model allows for evaluation of each case in terms of what RSD is feasible. The RSD at 0 is automatically infinite, no matter how accurate the measurement process is, and the RSD at high levels is σ_η so it is meaningful to define the quantification limit as the level at which the RSD falls to a specified factor of σ_η , say $2\sigma_\eta$. It makes little sense to specify a particular arbitrary RSD, such as 20%, since the limit of quantification would then be undefined for any measurement process with $\sigma_\eta > 0.20$. This is easily seen by substituting in 0.2 in the example given earlier. If it is desirable to make this computation for a particular process, this can easily be done using the model presented here. Let R be the desired *RSD*, e.g. 10%. Then the equation

$$q \frac{\text{---}}{V(L_Q)/L_Q} = R, \quad (4.16)$$

where $V(L_Q) = L_Q^2 S_\eta^2 + S_\epsilon^2$, has solution

$$L_Q = \frac{S_\epsilon}{\sqrt{R^2 - S_\eta^2}} \quad (4.17)$$

whenever $R > S_\eta$. No real solution is possible when $R \leq S_\eta$. This is readily apparent once (4.16) is re-written as a quadratic equation in L_Q .

For our zinc example, $S_\epsilon = 28.9$ ppt $S_\eta = .0390$. If the desired RSD is 10%, then

$$L_Q = 28.9 / \sqrt{(.1)^2 - (.0390)^2} = 314 \text{ppt}. \quad (4.18)$$

Compare this to the 99% confidence critical level of 67.2 ppt and the minimum detectable value of 135 ppt. The limit of quantification is somewhat arbitrary by comparison. While the critical level and the minimum detectable value are quantified using standard normal probability theory, the limit of quantification is set arbitrarily by the investigator. For example, if the target RSD is chosen to be 15%, rather than 10% then the quantification limit is 200 ppt instead of 314 ppt. For analytes that are toxic at very low levels, this arbitrary choice may have rather severe consequences. As in the case of the minimum detectable value, the limit of quantification has no use in interpreting measurements that have already occurred. The estimated concentration along with a measure of the uncertainty of the measurement convey all of the necessary information.

4.4. Sample size calculations

When detection is crucial, it makes sense to take replicate samples at a site rather than one measurement. The measurement error associated with the mean of the replicates is much smaller than the error associated with one measurement—thus replication allows for more sensitive detection of the toxin. Therefore, sample size calculations become quite important.

According to our model, the measurements at true concentration $\mu = 0$ are normally distributed with standard deviation σ_ϵ . If the number of replicates is r , then any average of measured values greater than

$$D = \alpha + \frac{3\sigma_\epsilon}{\sqrt{r}} \quad (4.19)$$

is extremely unlikely to have come from a zero concentration sample. Similarly, so is any average of estimated concentrations greater than

$$D = \frac{3S_\epsilon}{\sqrt{r}}. \quad (4.20)$$

(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 multiple measurements at low levels, even if they are individually below the critical level, can still provide quantitative evidence of the concentration of a toxic substance. If the safe level is near or below

the critical level, then requiring adequate replication of the measurements reduces the effective critical level. That is, much smaller amounts can be measured with greater certainty. Quantitative recording of repeated measurements, even when they are below the individual observation critical level, will allow quantitative evidence to be gathered to estimate the true concentration at a site.

The question remains—How many replicates are needed? 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 determine whether the concentration of zinc is above some specific level, say 50 ppt. From previous experience, we estimate the standard deviation parameters to be $S_\epsilon = 28.9$ ppt and $S_\eta = 0.0390$. Suppose further that it is held to be important to detect a concentration of 80 ppt and determine it to be out of compliance. At that concentration, the standard deviation of an average of r replicates is

$$\begin{aligned} \frac{S_\epsilon^2 + \mu^2 S_\eta^2}{r} &= \frac{(28.9)^2 + (80)^2 (.0390)^2}{r} \\ &= \frac{845}{r} \\ &= \frac{29.1}{\sqrt{r}} \end{aligned} \tag{4.21}$$

Using a normal approximation, the distance between the criterion level 50 ppt and the concentration 80 ppt in standard deviation units is

$$(30/29.1)\sqrt{r} = 1.03\sqrt{r} \tag{4.22}$$

For the chance of detection to be 0.95, we require

$$1.03\sqrt{r} > 1.645$$

or $r > 2.55$. In this case, then, we need at least three replicates.

4.5. Uncertainty of a Single Measurement

The uncertainty of a single measurement is usually quantified using confidence intervals. There are two primary approaches to this problem, an exact solution and a normal or lognormal approximation. The exact solution requires numerical integration and will not be discussed here. Say we would like a 95% confidence interval for μ based on a single measurement, $\hat{\mu}$. The approximate method for low values of $\hat{\mu}$, using an estimated variance and a normal approximation is

$$\hat{\mu} \pm 1.96 \sqrt{\text{Var}(\hat{\mu})} \tag{4.23}$$

where $\text{Var}(\hat{\mu})$ is estimated using

$$\text{Var}(\hat{\mu}) = \frac{\hat{\sigma}_\eta^2}{\hat{\beta}^2} + \hat{\mu} e^{\hat{\sigma}_\eta^2} (e^{\hat{\sigma}_\eta^2} - 1) \quad (4.24)$$

and where all estimates are obtained from the maximum likelihood routines or other methods. For high levels of $\hat{\mu}$, (those in which the second term in equation (2.5) dominates), $\ln \hat{\mu}$ is approximately normally distributed with variance σ_η^2 . Hence, a 95% confidence interval for μ is

$$(\exp(\ln \hat{\mu} - 1.96\hat{\sigma}_\eta), \exp(\ln \hat{\mu} + 1.96\hat{\sigma}_\eta)). \quad (4.25)$$

Note that this interval is symmetric on the log scale, but asymmetric on the original measurement scale.

For our zinc data, a measured concentration of 80.0 ppt would use the first technique and have a 95% confidence interval of $80.0 \pm \frac{(1.96)(29.1)}{\sqrt{(29.9)^2 + (80)^2(.0390)^2}} = 80 \pm 57.0$ (since the standard deviation of the 80.0 ppt measured value is $\sqrt{(29.9)^2 + (80)^2(.0390)^2}$). A measured value of 5000ppt would use the second method. The log measurement is 8.517 with standard deviation $\sigma_\epsilon = .0390$, so a 95% confidence interval on the log scale is $8.517 \pm (1.96)(.0390) = 8.517 \pm .076$, or (8.441, 8.593). Transforming the limits by exponentiation, we arrive at the 95% confidence limit (4630, 5390) for the concentration.

We can also use this method to give confidence intervals for the average of a series of replicate measurements. For low levels, the average of r measurements will be approximately normally distributed with standard deviation $\sqrt{\text{Var}(\hat{\mu})/r}$. For larger values of $\hat{\mu}$, the average of the natural log of the r measurements will have approximate standard deviation $\sigma_\epsilon/\sqrt{r}$. Confidence intervals can then be constructed as above, using the appropriate standard deviations.

5. Examples

In this section we present two examples of analytical methods and examine the fit of the two-component error model. Our first example is a metal data set using ICPMS. Our second example is an organic data set using GC/MS. In both cases, we use the calibration curve model (2.1)).

5.1. Zinc by ICPMS

This first example is a set of zinc data at 12 concentrations measured via inductively coupled plasma mass spectrometer (ICPMS), with 7 to 11 replications at each concentration. The maximum likelihood parameter estimates for α , β , σ_ϵ and σ_η are shown in Table 1. The summary data are shown in Table 2. The observed and predicted responses are shown in Figure 1. A plot of the residuals versus predicted values is shown in Figure 2. Note the increasing deviation of the observed response away from the calibration curve (i.e., the

predicted response). Thus, the data exhibit the error structure assumed by the model. The observed and model predicted variances are shown in Figure 3. Here also, the linear increase in standard deviation at high levels is clear, as well as an area at concentrations near zero where the standard deviation appears to remain somewhat constant.

5.2. Propionitrile by GC/MS

Our second example is a set of propionitrile data measured at 9 concentrations, using gas chromatography/mass spectrometry (GS/MS) with the number of replications varying between 4 and 11. The maximum likelihood estimation results are shown in Table 3. The summary data are shown in Table 4. The observed and predicted responses are shown in Figure 4. The residual versus predicted plot appears as Figure 5. The observed and predicted variances are shown in Figure 6. Once again, it can be easily observed that the data exhibit the error structure of the model and the estimation routine produces plausible results.

5.3. Conclusion

The two-component error model is useful for many applications in the assessment of environmental data since it provides accurate estimates of error across the entire usable range of a measurement technology, so long as the data exhibit the error structure specified by the model. The model has been tested on a wide variety of data sets, two of which were shown here. The estimation routine produces highly accurate maximum likelihood estimates for the model parameters for each of the data sets tested.

The two-component error model is especially useful in the calculation of critical values and limits of detection based on standard probability theory and also allows calculation of limits of quantification. Thus, the model provides a solid analytical framework for making detection decisions and a superior alternative to previous methods for calculating the quantities mentioned above. That is, the model facilitates explicit evaluation of the efficacy of alternative values for RSD as a criterion of quantification.

Appendix

A. Closed Form Solution for IUPAC Minimum Detectable Value

Given a critical level L_C with associated confidence level δ , and a desired confidence level δ^0 for L_D , the IUPAC minimum detectable value is the true concentration L_D such that the measured response will exceed L_C at least a fraction $1 - \delta^0$ of the time. The exact solution of this problem under the two-component model can be difficult, but a good approximation

can be derived using normal theory by letting L_D be the solution to the following equation

$$(\alpha + \beta L_D) - z_1 \frac{\mathbf{p} \overline{\text{Var}(y; L_D)}}{\mathbf{p} \overline{\text{Var}(y; L_D)}} = L_C = \alpha + z_0 \sigma_\epsilon \quad (1.26)$$

$$\beta L_D - z_1 \frac{\mathbf{p} \overline{\text{Var}(y; L_D)}}{\mathbf{p} \overline{\text{Var}(y; L_D)}} = z_0 \sigma_\epsilon \quad (1.27)$$

$$L_D - z_1 \frac{\mathbf{p} \overline{V(\hat{\mu}; L_D)}}{\mathbf{p} \overline{V(\hat{\mu}; L_D)}} = z_0 S_\epsilon \quad (1.28)$$

where $z_0 = z_\delta$ and $z_1 = z_{\delta^*}$, the appropriate quantiles of the normal distribution.

The equation

$$L_D - z_1 \frac{\mathbf{q} \overline{L_D^2 S_\eta^2 + S_\epsilon^2}}{L_D^2 S_\eta^2 + S_\epsilon^2} = z_0 S_\epsilon \quad (1.29)$$

has a solution for L_D exactly when

$$S_\eta < 1/z_1 \quad (1.30)$$

The solution is given by

$$L_D = \frac{\mathbf{h} \overline{S_\epsilon} + \mathbf{q} \overline{z_0^2 - (1 - z_1^2 S_\eta^2)(z_0^2 - z_1^2)}}{(1 - z_1^2 S_\eta^2)}. \quad (1.31)$$

If $z_0 = z_1 = z$, the solution has the particularly simple form

$$L_D = \frac{2zS_\epsilon}{(1 - z^2 S_\eta^2)}. \quad (1.32)$$

It is almost immediately apparent that L_D need not exist under the two-component model, or any model that allows the error variance to increase with the concentration. If $z_1 \frac{\mathbf{p} \overline{V(\mu)}}{\mathbf{p} \overline{V(\mu)}}$ exceeds μ for every value of μ , then the equation (1.28) certainly cannot be solved. This in turn will be true if

$$z_1^2 e^{\sigma_\eta^2} (e^{\sigma_\eta^2} - 1) > 1 \quad (1.33)$$

$$z_1^2 S_\eta^2 > 1 \quad (1.34)$$

$$S_\eta > 1/z_1 \quad (1.35)$$

Thus, it is necessary for L_D to exist that $S_\eta < 1/z_1$, and it also turns out to be sufficient.

We can solve the defining equation for L_D in the following way

$$-z_1 \frac{\mathbf{q} \overline{L_D^2 S_\eta^2 + S_\epsilon^2}}{L_D^2 S_\eta^2 + S_\epsilon^2} = z_0 S_\epsilon - L_D \quad (1.36)$$

$$z_1^2 L_D^2 S_\eta^2 + z_1^2 S_\epsilon^2 = (z_0 S_\epsilon - L_D)^2 \quad (1.37)$$

$$z_1^2 L_D^2 S_\eta^2 + z_1^2 S_\epsilon^2 = z_0^2 S_\epsilon^2 - 2z_0 S_\epsilon L_D + L_D^2, \quad (1.38)$$

or

$$L_D^2 (1 - z_1^2 S_\eta^2) - L_D (2z_0 S_\epsilon) + (z_0^2 - z_1^2) S_\epsilon^2 = 0. \quad (1.39)$$

It is not difficult to show that this quadratic equation will have real solutions whenever (1.30) holds. The discriminant of (1.39) is

$$D = 4z_0^2 S_\epsilon^2 - 4(1 - z_1^2 S_\eta^2)(z_0^2 - z_1^2) S_\epsilon^2 \quad (1.40)$$

$$= 4S_\epsilon^2 [z_0^2 - (1 - z_1^2 S_\eta^2)(z_0^2 - z_1^2)]. \quad (1.41)$$

This can only be negative if the second term within square brackets is negative (since the first term is positive). By hypothesis, $1 - z_1^2 S_\eta^2 > 0$, so that the only way this term can be negative is when $z_0^2 > z_1^2$. However, $0 < 1 - z_1^2 S_\eta^2 < 1$, so under these conditions,

$$z_0^2 - (1 - z_1^2 S_\eta^2)(z_0^2 - z_1^2) > z_0^2 - (z_0^2 - z_1^2) = z_1^2 > 0 \quad (1.42)$$

Hence, the discriminant is always positive when condition (1.30) is satisfied.

The appropriate solution to this quadratic equation satisfies

$$L_D = \frac{(2z_0 S_\epsilon) + 2S_\epsilon \sqrt{z_0^2 - (1 - z_1^2 S_\eta^2)(z_0^2 - z_1^2)}}{2(1 - z_1^2 S_\eta^2)} \quad (1.43)$$

$$= \frac{S_\epsilon z_0 + \sqrt{z_0^2 - (1 - z_1^2 S_\eta^2)(z_0^2 - z_1^2)}}{(1 - z_1^2 S_\eta^2)}, \quad (1.44)$$

and this solution is positive exactly under our hypothesis. If $z_0 = z_1$, the solution has the particularly simple form

$$L_D = S_\epsilon \frac{2z_0}{(1 - z_1^2 S_\eta^2)}. \quad (1.45)$$

Intuitively, the limit of detection would be $S_\epsilon(z_0 + z_1)$ if the variance were constant, and the factor in the denominator inflates the result to account for the increasing variance.

Let us consider an example. Suppose that $\alpha = 0$ and $\beta = 1$, so that the concentration and the response are on the same scale, and suppose that the measurement process parameters are $\sigma_\epsilon = S_\epsilon = 1$ and $\sigma_\eta = 0.1$. Then

$$S_\eta = \frac{1}{e^{\sigma_\eta^2}(e^{\sigma_\eta^2} - 1)} = .10075, \quad (1.46)$$

which is essentially the same as σ_η . When the confidence level for both the critical value and the detection limit are chosen to be 95%, that is $z_0 = z_1 = 1.645$, application of the formulas leads to $L_C = 1.645$ and $L_D = 3.383$ (only slightly greater than if the variance did not increase with the concentration). If a confidence level of 99% for both is desired then, $z_0 = z_1 = 2.326$ and $L_C = 2.326$ and $L_D = 4.923$.

On the other hand, if $\sigma_\eta = .3$, then $L_C = 2.326$ as before, but L_D now increases to 10.518. Finally, when $\sigma_\eta = .385$, we cannot find a solution for L_D since $S_\eta = .4305 > 1/2.326 = .4299$. When the variance increases this rapidly with the concentration, no concentration can guarantee at the 99% confidence level that the measured value will exceed L_C .

Acknowledgments

Research reported in this paper was supported by grants from the United States Environmental Protection Agency (CR 825621-01-0) and the National Institute of Environmental Health Sciences (P42 ES 04699). The data used as examples are courtesy of the National Council of the Paper Industry for Air and Stream Improvement (organics) and the US EPA (metals).

References

- Cohen, A.C. (1991) *Truncated and Censored Samples*, New York, Marcel Dekker.
- Currie, L. (1968) "Limits for qualitative detection and quantitative determination—Application to radiochemistry," *Analytical Chemistry*, Vol. 40, No.3, p.586
- Currie, L. (1995) "Nomenclature in Evaluation of Analytical Methods Including Detection and Quantification Capabilities," *Pure & Applied Chemistry*, Vol. 67, No. 10, pp.1699–1723.
- Currie, L.A. (1997) "Detection: International Update, and some emerging dilemmas involving calibration, the blank, and multiple detection decisions." *Chemometrics and Intelligent Laboratory Systems*. Vol. 37, No. 1, pp. 151–181.
- Gibbons, R. (1994) *Statistical Methods for Groundwater Monitoring*, Wiley.
- Gilbert, R.O. (1987) *Statistical Methods for Environmental Pollution Monitoring*, New York: Van Nostrand Reinhold.
- Glaser, J.; D.L. Foerst, G.D. McKee, S.A. Quave, and W.L. Budede, (1981) "Trace Analyses for Wastewaters," *Environmental Science and Technology*, Vol. 15, No. 12, pp.1426–1435.
- Hubaux, A. and Vos, G. (1970) "Decision and detection for linear calibration curves," *Analytical Chemistry*, Vol. 42, No. 8, pp.849–855.
- Liteanu, C. and Rica, I. (1980) *Statistical Theory and Methodology of Trace Analysis*, Chichester, UK: Ellis Horwood.
- Meier, P.C. and Zünd, R.E. (1993) *Statistical Methods in Analytical Chemistry*, Wiley.
- Oppenheimer, L.; T.P. Capizzi, R.M. Weppelman, and H. Mehta (1983) "Determining the Lowest Limit of Reliable Assay Measurement," *Analytical Chemistry*, Vol. 55, No. 4, pp.638–643.
- Rocke, D.M., S. Lorenzato; (1995) "A Two-Component Model for Measurement Error in Analytical Chemistry," *Technometrics*, Vol. 37, No. 2, pp.176–184.

- Zorn, M.E., Gibbons, R.D., and Sonzogni, W.C. (1997) "Weighted Least-Squares Approach to Calculating Limits of Detection and Quantification by Modeling Variability as a Function of Concentration," *Analytical Chemistry*, Vol. 69, 3069–3075.
- Zorn, M.E., Gibbons, R.D., and Sonzogni, W.C. (1999) "Evaluation of Approximate Methods for Calculating the Limit of Detection and the Limit of Quantification," *Environmental Science and Technology*, Vol. 33, 2291–2295.

Table 1: Zinc Parameter Estimates

Parameter	Maximum Likelihood Estimate
α	490
β	7.06
σ_{ϵ}	204
σ_{η}	0.0390
S_{ϵ}	28.9ppt
S_{η}	0.0390
L_C (peak area)	965
L_C (concentration)	67.1ppt
L_D	135ppt
L_Q	314ppt

The MLEs were obtained using numerical methods in Fortran and then used to calculate the critical value, L_C , limit of detection, L_D , and limit of quantification, L_Q . The first value reported for L_C is peak area; the second is in units concentration (parts per trillion). The L_D and the L_Q are both in units of concentration. L_Q was calculated using a CV of 0.10. Both L_D and L_C were calculated using a 99% confidence level.

Table 2: Zinc data Set

Concentration in parts per thousand	Number of reps	Average Observed Response (peak area)	Predicted response
0	8	265	490
10	7	317	561
20	7	692	631
100	11	1,290	1,200
200	7	2,190	1,900
500	7	4,110	4,020
1000	9	7,590	7,550
2000	7	14,400	14,600
5000	9	35,100	35,800
10000	10	70,300	71,100
25000	9	181,000	177,000

The predicted response at each concentration was calculated using the maximum likelihood estimates for the parameters of the calibration curve.

Table 3: Propionitrile Parameter Estimates

Parameter	Maximum Likelihood Estimates
α	559
β	18.7
σ_{ϵ}	147
σ_{η}	0.0397
S_{ϵ}	7.85ppb
S_{η}	0.0398
L_C (peak area)	900
L_C (concentration)	18.3ppb
L_D	36.8ppb
L_Q	85.6ppb

The MLEs were obtained using numerical methods in Fortran and then used to calculate the critical value, L_C , limit of detection, L_D , and limit of quantification, L_Q . The first value reported for L_C is peak area. The second is in units concentration (parts per billion). The L_D and the L_Q are both in units of concentration. L_Q was calculated using a CV of 0.10. Both L_D and L_C were calculated using a 99% confidence level.

Table 4: Propionitrile Data Set

Concentration in parts per billion	Number of reps	Average Observed Response (peak area)	Predicted response
4.8	7	7.16	7.37
16.0	8	18.1	17.4
32.1	7	31.7	31.7
48.1	3	42.0	46.0
160	6	151	146
321	4	284	289
481	3	431	432
722	5	658	647
3,010	4	2,660	2,680

The predicted response at each concentration was calculated using the maximum likelihood estimates for the parameters of the calibration curve.

Figure 1: Zinc
Predicted and Observed Responses

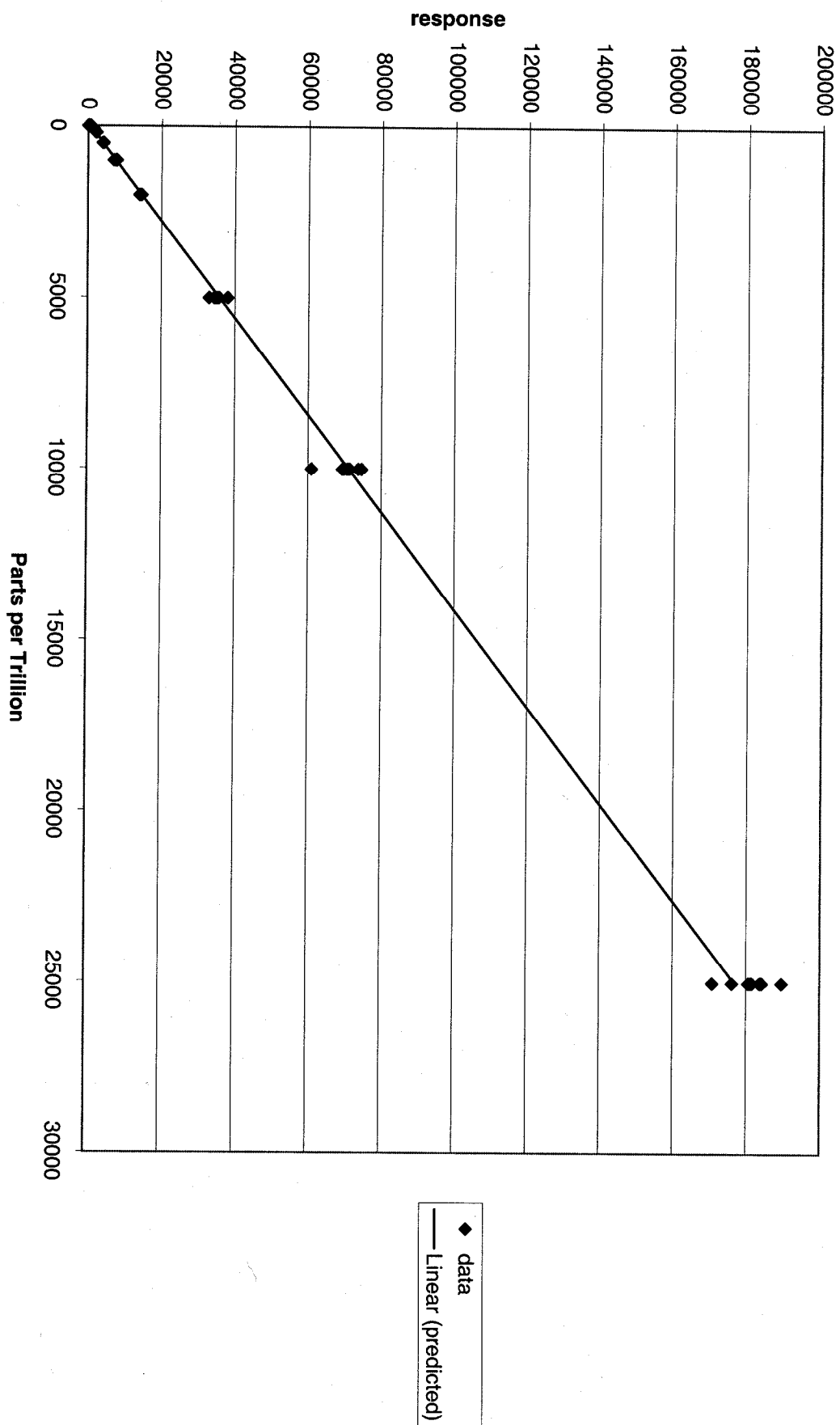


Figure 2: Zinc Residuals

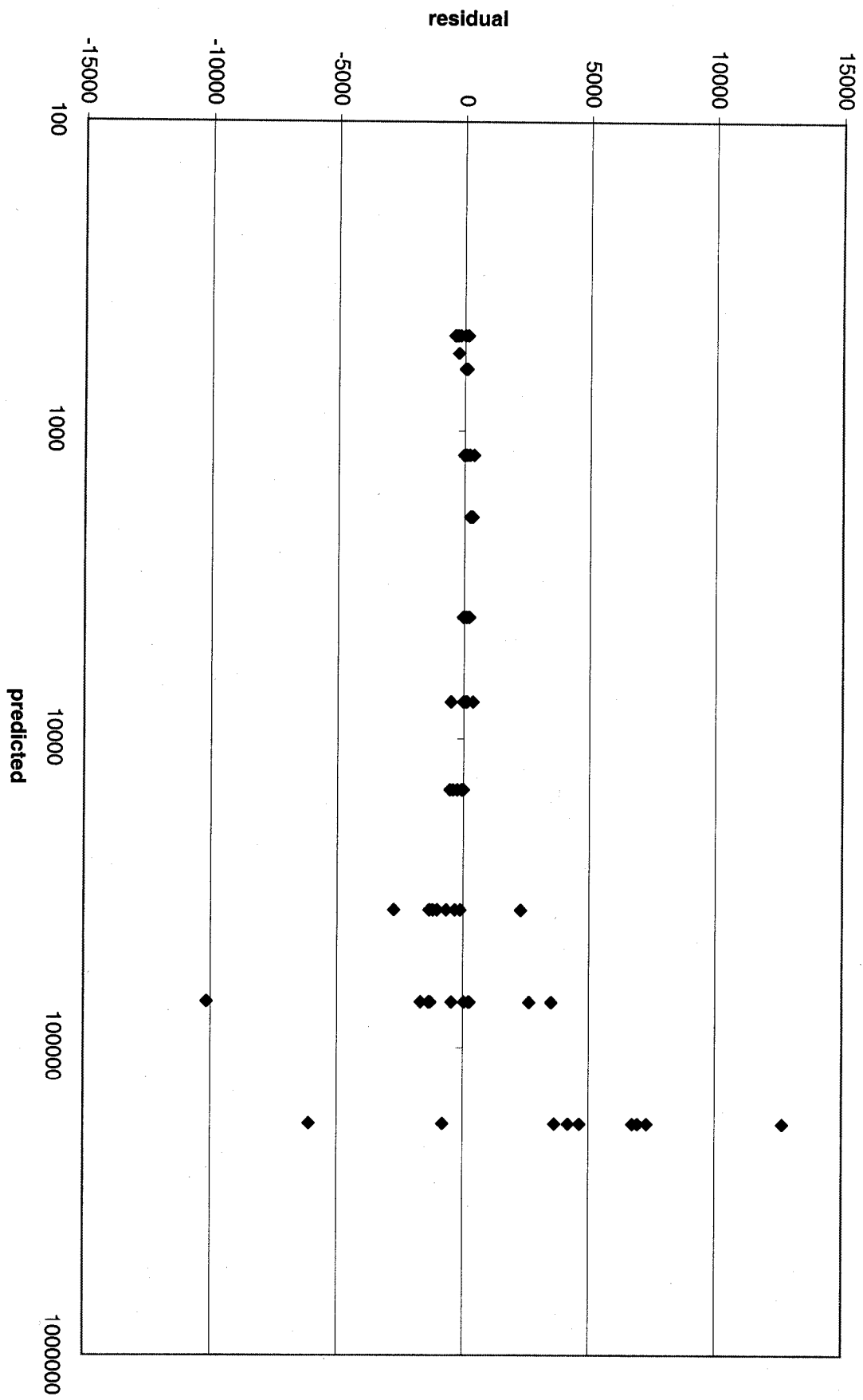


Figure 3:
Predicted vs Observed Standard Deviation for Zinc

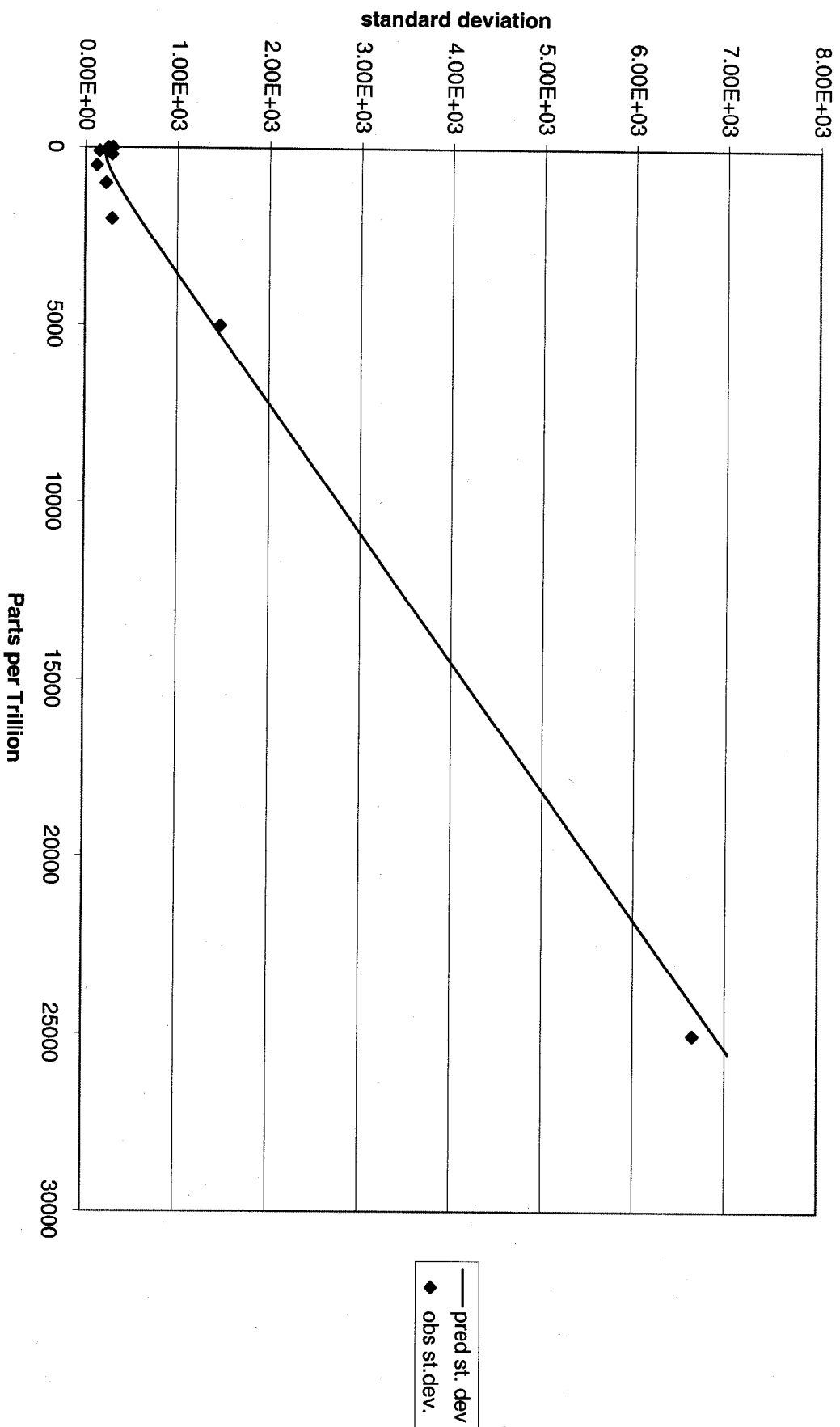


Figure 4: Propionitrile
Observed and Predicted Responses

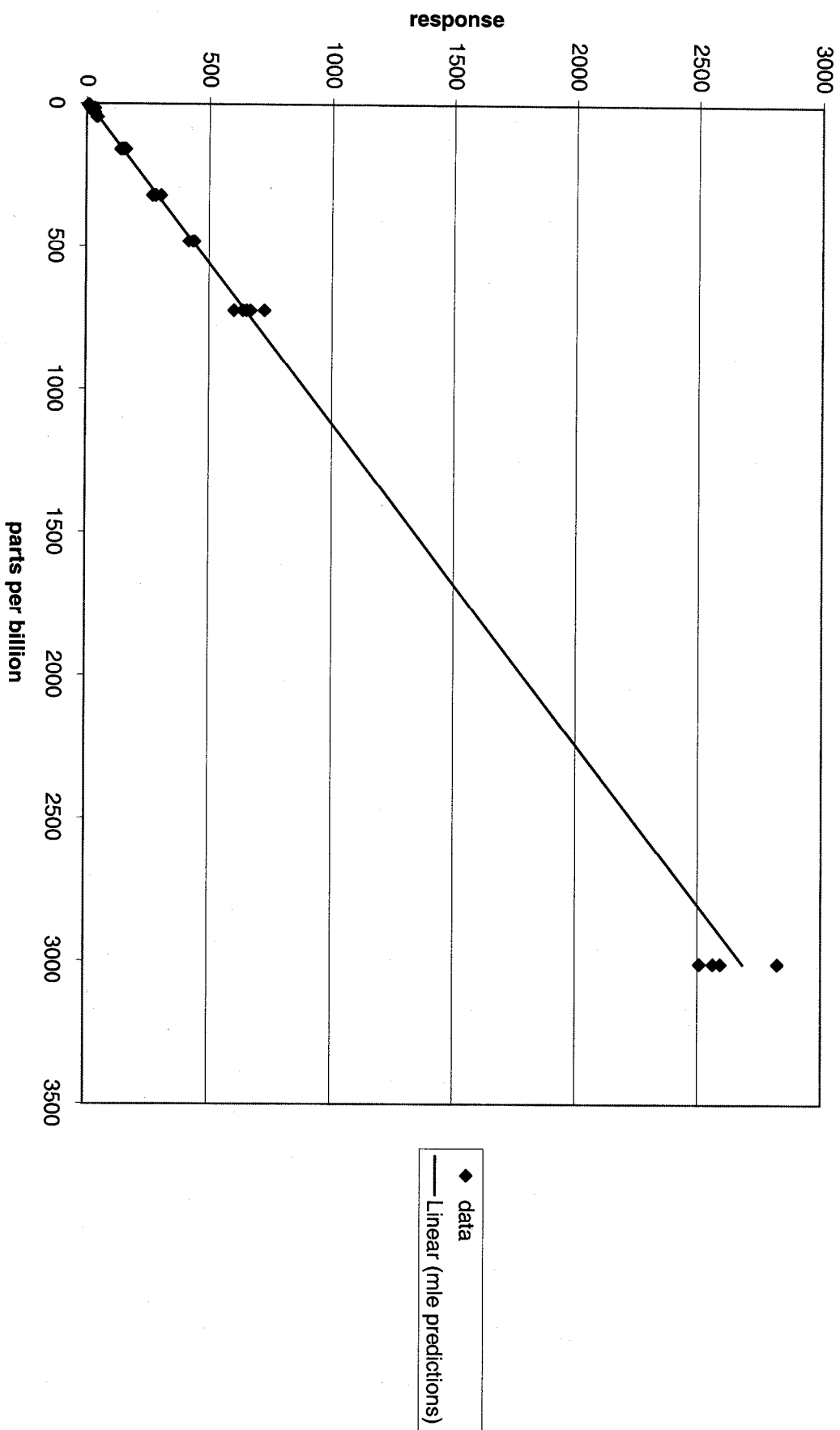


Figure 5: Propionitrile residuals

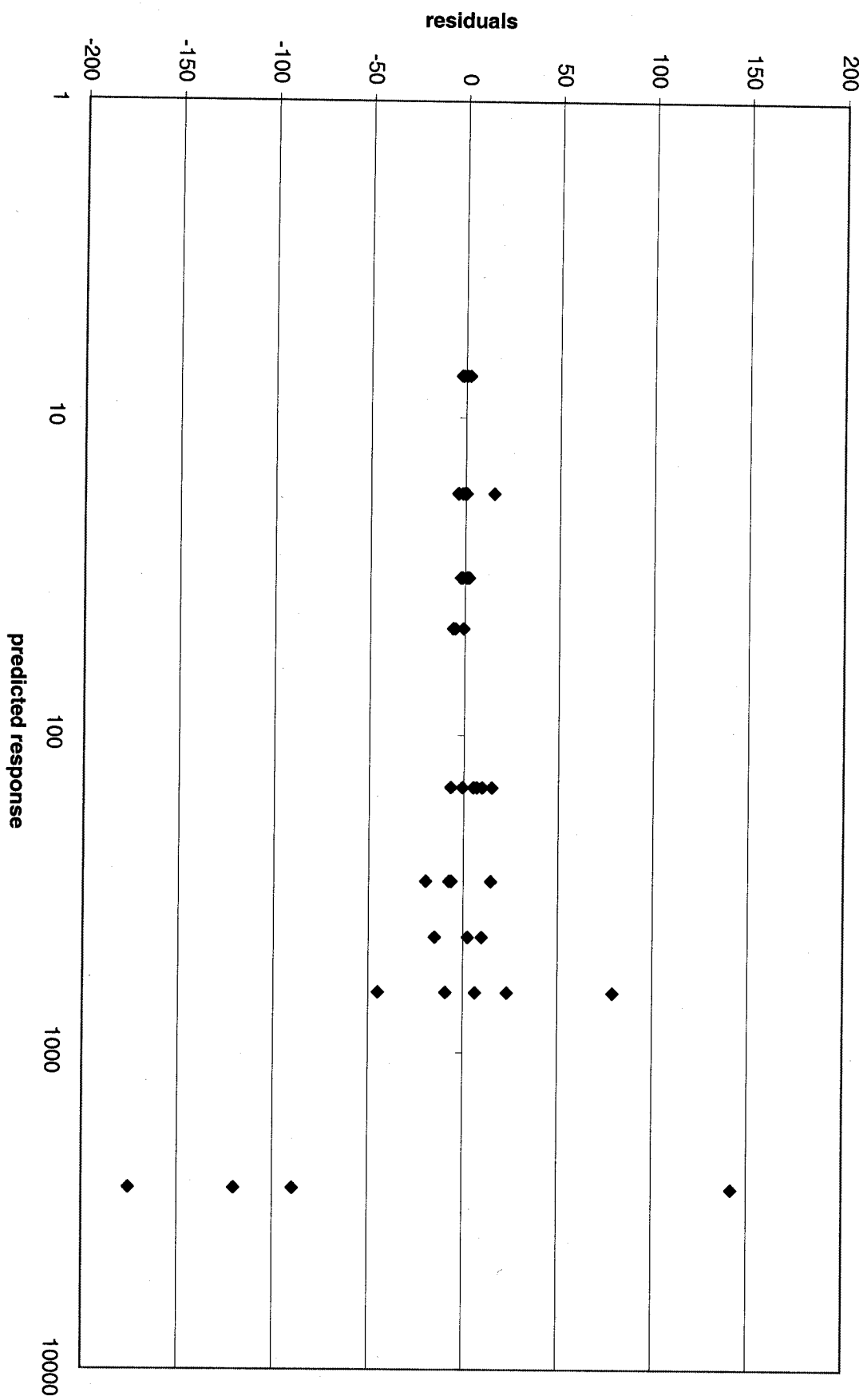


Figure 6:
Predicted vs Observed Standard Deviation for Propionitrile

