REPEATABILITY (PRECISION) OF DIOXIN STACK GAS CONCENTRATION MEASUREMENTS

H. Gregor Rigo

17168 Misty Lake Drive, Strongsville, Ohio, 44136-7357, USA

Introduction

Repeatability (precision) of a measurement technique describes the range of likely results if another simultaneous measurement had been made. The repeatability of an environmental measurement is important because regulators and the public need to know if a source is in or out of compliance. Sources and suppliers need to know how much margin to provided between the design and compliance points so a facility will pass whenever it is tested. Finally, researchers need to know if standard statistical models that assume the predictors are known much more precisely than the responses can be used or if special techniques that do not have this restriction apply.

The focus has historically been on determining the detection and quantification limits for emitted concentration measurement systems[1]. Current thought, however, is that the relevant question is the repeatability at any level of interest, because existing dioxin and POP measurement methods and results are all that we have. In any case, once repeatability is known, then appropriate compliance, design and scientific decisions can be made. Incorrect conclusions and consequent decisions are likely when repeatability is either unknown or not properly considered.

Methods and Materials

For complete sampling and analysis methods, repeatability is determined by studying the differences between collocated, simultaneous samplings. Unlike a chemist who can make up a standard solution and generate a virtually unlimited number of replicate analyses, only two to four simultaneous samples can be obtained with practical dual- and quad-sampling trains. These sampling trains extract samples either within 2.5 cm of each other or from a 6 cm square. This closeness exhibits minimal spatial variability[2]. For small samples, however, the calculated standard deviation converges to some fraction of the population standard deviation[3]. The sample standard deviation is multiplied by the small sample bias correction factor (1.253, 1.128 and 1.085 for pairs, triplets and quads respectively) to estimate the population standard deviation[4]. This oft overlooked correction is particularly important when combining simultaneous test results based on different numbers of replicates.

Simultaneous samplings are sometimes taken at different locations along a branch-free duct, but spatial changes induced by gravitational settling, electrostatic and flow effects cannot be ruled out. These possibilities must be considered in the analysis.

When the variability of the response variable changes with concentration, and most emitted concentration measurements exhibit this characteristic, an implicit assumption in regression analysis is violated. The residual differences between the predicted and measured dependent variables must be independent of changes in the predicted values. A practical approach is to work with the logarithms of the concentrations and bias-corrected standard deviations. Other transformations may have to be employed if the residuals change or display pronounced curvature. When nonlinear transformations of the response variable are used, the average of the standard

deviations after retransforming the predicted values will not equal the average of the original untransformed values. Consequently, regression results are multiplied by ratio of the original to retransformed standard deviation averages to eliminate retransformation bias[5].

Weighted least squares regression must be used to analyze the data when different numbers of samples are used in the standard deviation estimates. The weights are the degrees of freedom associated with each simultaneous set of data. Ordinary least squares (OLS) can be used by repeating the each of the N concentration-standard deviation data pairs a number of times equal to its degrees of freedom (v). The regressions statistics generated by the OLS program are all correct except for the standard error of the regression, which must be multiplied by $\sqrt{(N-1)/(\sum v - 1)}$.

Because the standard deviation – concentration relationship is based on a sample, there is no way to tell if the realization is nearly right or some distance from the real answer. This difficulty is handled by using the upper confidence limit for the regression line[6] to estimate the population standard deviation at any concentration of interest. Many textbooks provide equations for determining the confidence limit for a point on a line. The difference between those equations and the confidence limit for the regression line is the use of $\sqrt{2 F_{2,N-2,1-\alpha/2}}$ for the line instead of

$t_{v,\alpha/2}$ for a single point on the regression line (α is the statistical significance used in the F- and t-statistics). Complete formulas can be found elsewhere[7, 8]. To understand the need to use the line rather than a point on the line, just consider regulating different sources using a common diluent corrected standard – facilities operating at 3, 7 or 10 percent excess oxygen will all have different uncorrected concentrations when they are emitting the same amount of diluent corrected dioxin.

There is no need to begin with diluent corrected results, but the data must all be expressed at the same temperature, pressure and moisture content. Monte Carlo simulations demonstrate that simply multiplying the repeatability estimate by the dilution correction factor produces a result that is within 1 percent of the true value the vast majority of the time. The largest deviation (4 percent) occurred when a number of unusually severe assumptions were realized simultaneously. The conversion to other temperature, pressure or moisture conditions similarly affects the concentration and repeatability and can be handled the same way.

As a final step, the upper confidence limit for the population standard deviation needs to be multiplied by the number of standard deviations above the mean likely to include a specified percentage of the realized results. The multiplier is 1.960 for 95% inclusion and 2.576 for 99% inclusion in this analysis. Multipliers of 2 and 3, which correspond 95.4 and 99.7 percent inclusion respectively, are frequently used for simplicity[9].

Results and Discussion
There are three sets of simultaneous dioxin sampling data available: the Pittsfield data was generated using dual-trains at a small ESP equipped incinerator during an experiment that created a range of dioxin concentrations using powdered activated carbon injection; the lightweight aggregate kiln (LWAK) data was collected simultaneously by two teams sampling on opposing traverses; the EN-1984 data comes from simultaneous samplings taken along a duct using three different sampling methods, two analytical laboratories and several teams at three different plants[10, 11, 12]. The outlier identified in the EN-1948 data set was confirmed and excluded from this analysis. Statistical tests using dummy variables to isolate specific effects verified that LWAK

cross-traverse data and the valid EN-1984 sampling data from all methods are comparable. Of course, additional data might alter this conclusion. All three data sets were used to determine the repeatability of the ITEQ dioxin measurements; only the Pittsfield and LWAK data sets include the homologue totals needed to determine the repeatability of Total dioxin measurements. Figure 1 is a graphical summary of the interval likely to contain 99 percent of the individual measurements (95% statistical confidence level). The limits narrow when m runs are averaged; the bands collapse by $\sqrt{m}$. The statistical characteristics of the curve fit and underlying data are summarized in Table 1.

Figure 1. Repeatability (precision) of stack ITEQ dioxin concentration measurements over the data range.
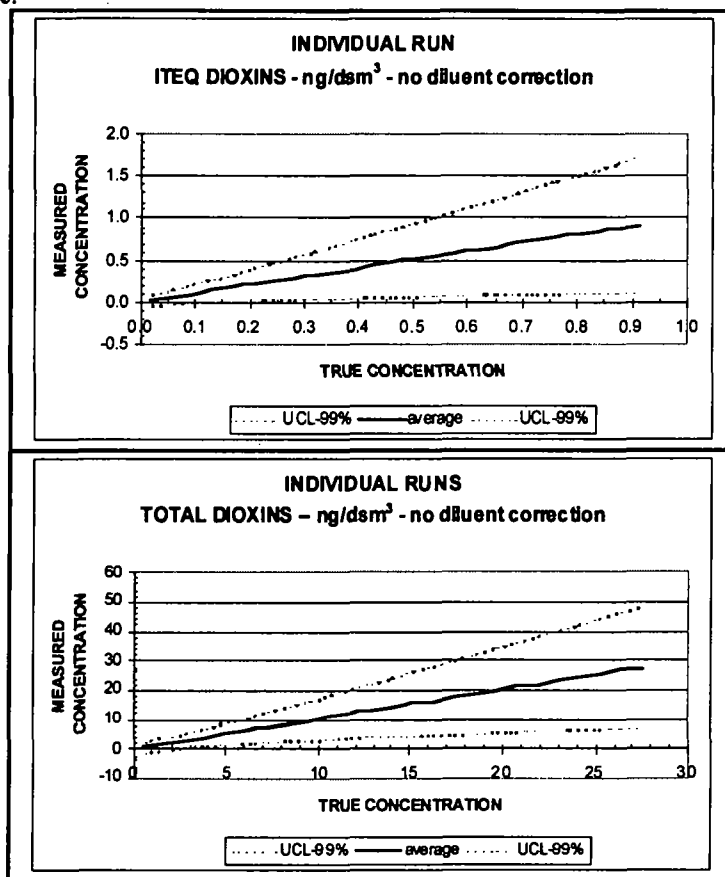
The intercepts are different for ITEQ and Total Dioxins. This is expected since the ITEQ concentrations associated with a given Total Dioxin measurement is usually between 1/40th and 1/100th. The slopes are the statistically the same which indicates that the congener and homologue distributions do not change for this data set over the data range. The repeatability graphics indicate that over the data range, it is very difficult to determine that any ITEQ result is substantially above zero, but it is relatively clear that the maximum likely concentrations are below some threshold.

Table 1. Summary statistics based on the available ITEQ and Total Dioxin Homologue data.

	ITEQ Dioxins		Total Dioxin Homologues	
	Value	Std. Error	Value	Std. Error
intercept	-3.228	0.564	-1.939	0.372
slope	0.492	0.211	0.559	0.23
standard error of the estimate	1.4324		1.2673	
effective number of samples N	46		22	
average of ln(C)	-2.4789		1.1089	
standard deviation of ln(C)	0.9578		1.2018	
retransformation bias correction factor	1.922		1.894	

References

1. Rigo H. G. and A. J. Chandler, *How fuzzy are point source dioxin measurements (USEPA Method 23)*, in: R. Hites, ed., Organohalogen Compounds, Volume 31, Indianapolis, August 25-29, 1997, pp 154-9.

2. Mitchell, W. J. and M. R. Midgett, *Means to evaluate performance of stationary source test methods*, ES&T, V10, N1, January 1976, pp 85-88.

3. Natrella, M. G., Experimental Statistics, NBS Handbook 91, USGPO, 1966, pp 1-10.

4. Dixon, W. J., and F. J. Massey, Introduction to Statistical Analysis, 3d ed., McGraw-Hill, New York, 1969, p 136.

5. Duan, N., *Smearing estimate: a nonparametric retransformation method*, JASA, V78, N383, September 1983, pp 605-10.

6. Natrella, M. G., Experimental Statistics, NBS Handbook 91, USGPO, 1966, pp 5-13 thru –27.

7. Lanier, W. S., and C. D. Hendrix, Reference method accuracy and precision (ReMAP) -- Phase I – precision manual air emission measurements, ASME, in process.

8. Rigo, H. G., *Dioxin Measurement Method Precision*, Proceedings of the North American Waste-to-Energy Conference VIII, SWANA, Silver Springs, MD, 2000.

9. ASME, Test Uncertainty – Instruments and Apparatus – Supplement to ASME Performance Test Codes, ASME, PTC-19.1-1998.

10. Rigo H. G. and A. J. Chandler, Retrofit of waste-to-energy facilities equipped with electrostatic precipitators – sensible heat exchanger cooling – Phase II, ASME, CRTD-Vol. 44, 1997.

11. Energy and Environmental Research Corporation (EER), Dioxins/Furans, HCl, Cl2, and related testing at a hazardous waste burning lightweight aggregate kiln, USEPA contract No. 68-D2-0164, October 10, 1997.

12. Comite Europeen de Normalisation, Stationary source emissions – determination of the mass concentration of PCDDs/PCDFs – Part 3: Identification and quantification, EN 1948-3, December 1996.