A Robust Approach for the Determination of CCQM Key Comparison Reference Values and Uncertainties

David L. Duewer

Analytical Chemistry Division
Chemical Science and Technology Laboratory
National Institute of Standards and Technology
Gaithersburg, Maryland 20899-8390
david.duewer@nist.gov

On 14 October 1999, the directors of the national metrology institutes (NMIs) of thirty-eight Member States of the Metre Convention and representatives of two international organizations signed a Mutual Recognition Arrangement (MRA) for national measurement standards and for calibration and measurement certificates issued by NMIs. This MRA was a response to a growing need for an open, transparent and comprehensive scheme to give users reliable quantitative information on the comparability of national metrology services and to provide the technical basis for wider agreements negotiated for international trade, commerce and regulatory affairs.

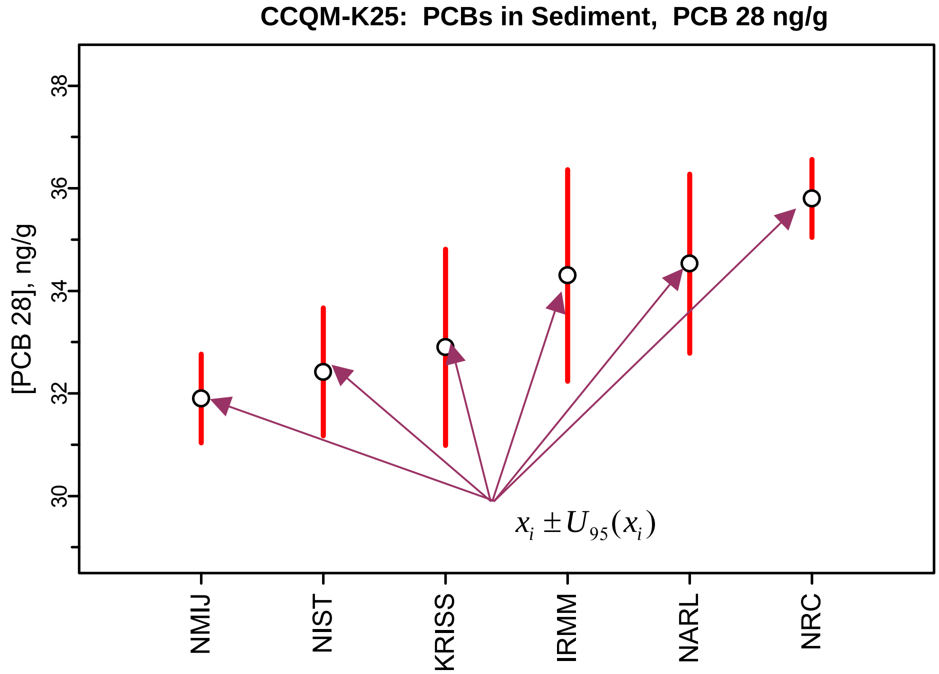
Signatories of the MRA committed themselves to active participation in Key Comparison (KC) studies conducted under the auspices of the Bureau International des Poids et Mesures (BIPM). While similar in design to standard proficiency test (PT) interlaboratory studies, the KCs are unique in several aspects:

- KC participants are the entire population of NMIs claiming competence in the particular type of measurement at a particular time.
- KC measurement results are identified by participant name.
- KC measurement results are required to be expressed as an expected value, x_i, within a fully evaluated 95% coverage uncertainty interval, $U_{95}(x_i)$.

The first of these considerations is not compatible with treating results as a random sample from an infinite population; the second tends to make participants much more critical of any data analysis; and the third enables definition of a meaningful mixture model probability density function (MM-PDF) for each measurement population.

The MM-PDFs provide easily interpreted visual summaries of the $\{x_i, U_{95}(x_i)\}$ measurement kernels that complement and extend the utility of the familiar "dot and bar" $x_i \pm U_{95}(x_i)$ plots (**Figure 1**). In addition to enabling visualization and discussion of metrologically interesting aspects of the data, MM-PDFs facilitate definition of robust estimates of location and dispersion that fully utilize all of the information provided by the kernels. These MM-PDF-based summary statistics are suitable for estimating 1) the expected performance of the majority of participants when some of the results are not in accord with the majority and 2) the true value of a measurand in a given material. Whether these or any other form of robust or trimmed statistics are suitable for estimating KC Reference Values (KCRVs) and their associated uncertainties, u(KCRV), is a philosophical – not a numerical – issue.

Figure 1: Dot & Bar ($x_i \pm U_{95}(x_i)$) Graph



What is a Mixture Model Probability Density?

Every measurement expressed as an expected value, x_i, and some defined uncertainty on x_i, $U(x_i)$, defines a probability density within which the true value of the measurand is expected to lie with specified confidence. While the shape of the density is seldom explicitly defined, in the absence of other information it is nearly universal practice to assume that the distribution is approximately normal (i.e., Gaussian) in form. Given that ±2 standard deviations about the mean include about 95% of the area of a normal distribution, KC measurements reported as $\{x_i, U_{95}(x_i)\}$ can be considered to specify that the true value of the measurand is normally distributed about x_i with standard deviation of $U_{95}(x_i)/2$. That is, each KC measurement reported as $\{x_i, U_{95}(x_i)\}$ specifies a normal kernel density of unit area: $N(x_i, U_{95}(x_i)/2)$ (**Figure 2**).

Displaying the individual kernels along the "bars" becomes ineffective (not to mention visually fussy) as the number of measurements displayed grows large. They are more usefully displayed along the margin of the plot; i.e., as a kind of marginal distribution (**Figure 3**).

Figure 2: Interpretation of $x_i \pm U_{95}(x_i)$ as N(x_i, $U_{95}(x_i)/2$) Kernel Density

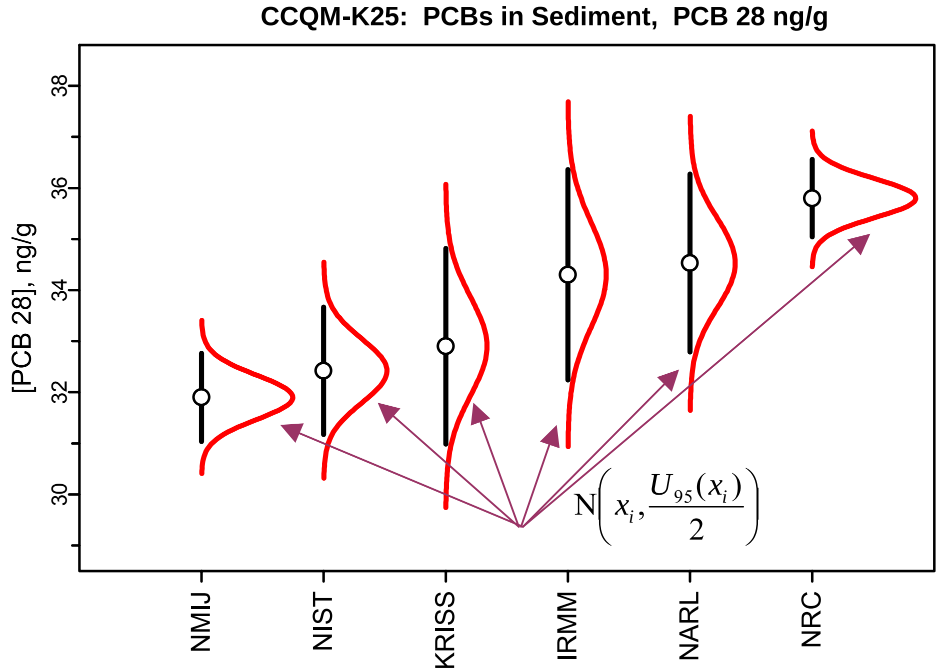
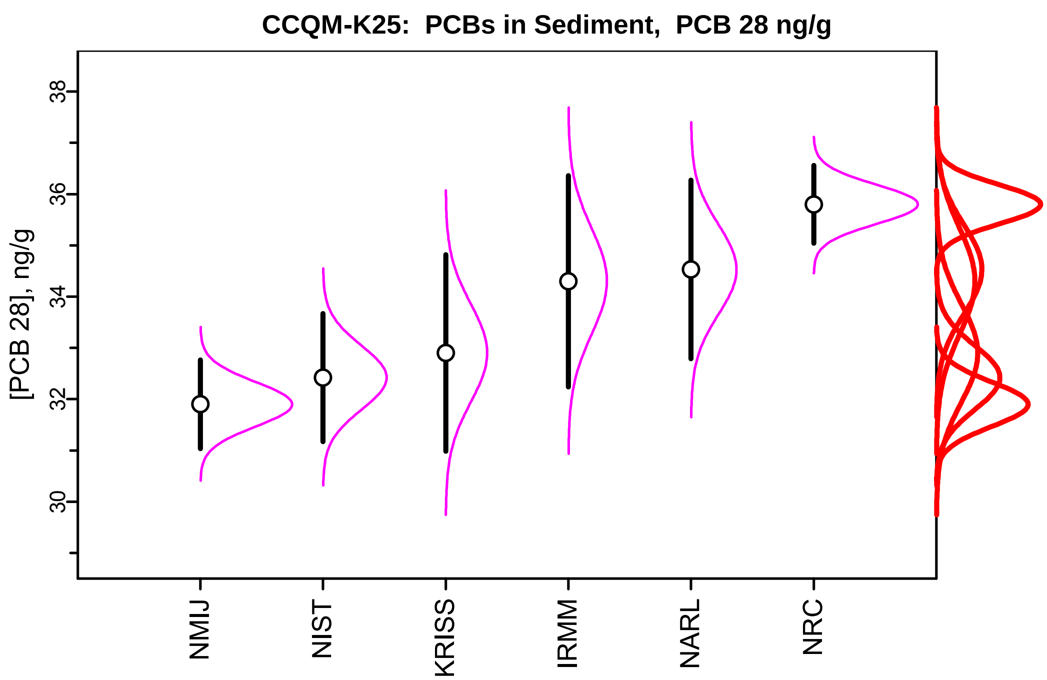
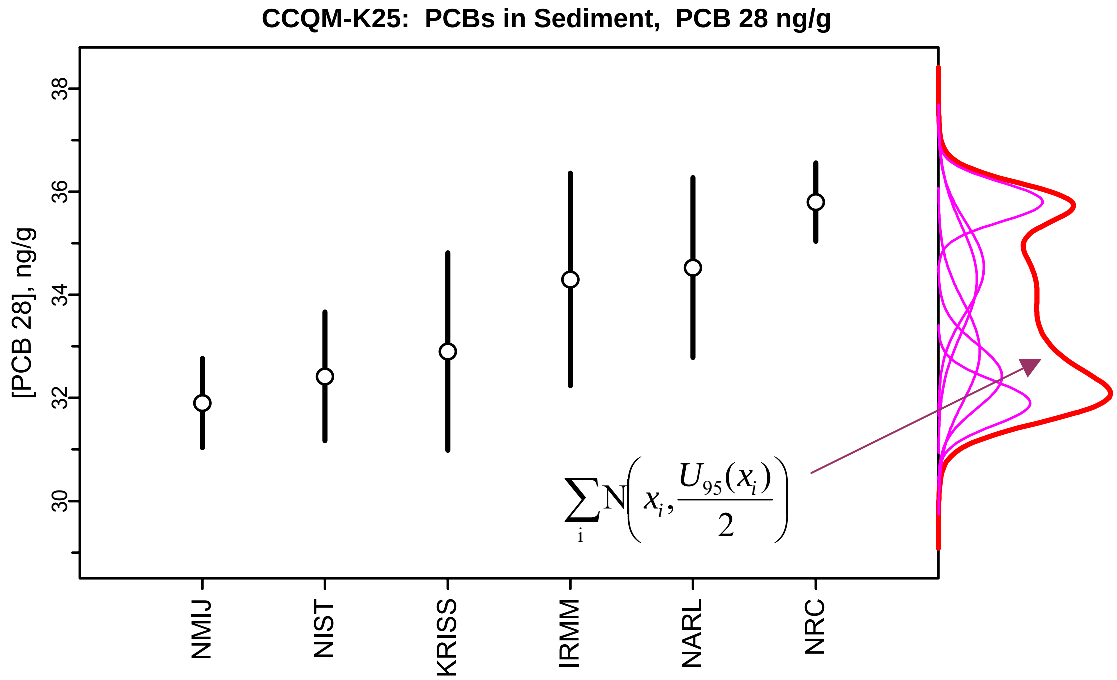


Figure 3: Visualization of Kernel Densities as Marginal Distributions



While the individual kernel densities summarize each of the measurements of a given KC, they do not themselves aggregate the measurements into a meaningful summary. However, this can be achieved by additively mixing the individual kernel densities together; i.e., the mixture model probability density function: MM-PDF = $\sum_{i} \mathrm{N}\left(x_i, \frac{U_{95}(x_i)}{2}\right)$ (**Figure 4**).

Figure 4: Interpretation of Kernel Densities as a MM-PDF



The MM-PDF is not a new tool for analysis of interlaboratory data. The conceptually related Kernel Density Estimation (KDE) techniques have long been used to provide smooth histograms and to define representative within-participant standard deviations for all values of given type reported in a given study [1]. KDE has also been advocated as an aid to detecting multimodalities within a dataset [2]. A computationally complex but explicit use of individual precision estimates for defining composite probability density functions has been used in the analysis of the results from several organic trace pollutant studies [3, 4]. Mixture models have recently been used with KC thermometry data [5, 6].

The MM-PDF is a useful interpretation tool in and of itself!

While the MM-PDF does not use any information that is not implicit in the familiar "dot and bar" plots, it explicitly summarizes that information. The weakly multiple-modes of the MM-PDF of the CCQM-K25 results (**Figure 4**) suggest that there may real metrological differences among the participants. But do these differences arise from the measurement processes (the x_i) or from the participant's evaluations of the possible measurement uncertainty components (the $U_{95}(x_i)$)?

Figure 5 displays the MM-PDF for the data assuming that all of the uncertainties are as large as the largest actually reported. If the participants reporting the smaller $U_{95}(x_i)$ have under-estimated their uncertainties, the nearly symmetrical single-mode shape of the MM-PDF would suggest the absence of significant differences among the participant's results. In contrast, **Figure 6** displays the MM-PDF for the data assuming that all of the uncertainties are as small as the smallest actually reported. If the participants reporting the larger $U_{95}(x_i)$ have over-estimated their uncertainties, the strongly multiple-modes would suggest that the participant's methods do significantly differ. No statistic can resolve the issue, but the MM-PDF can usefully guide metrological investigation into root causes.

Figure 5: If all of the $U_{95}(x_i)$ were as large as the largest reported value…

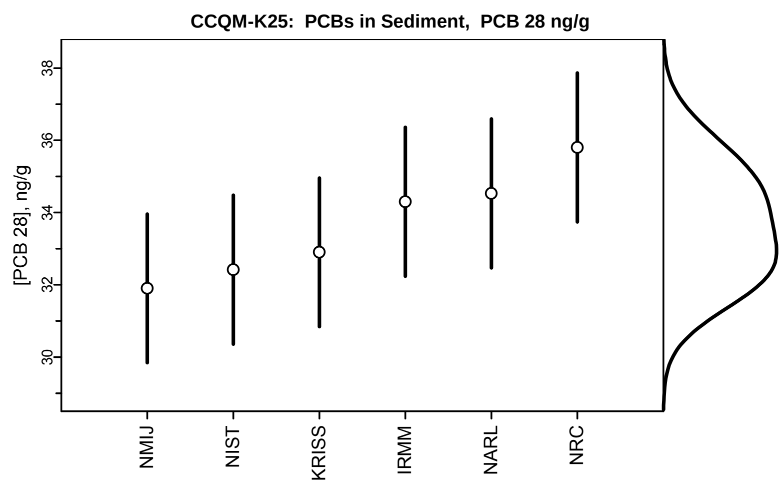
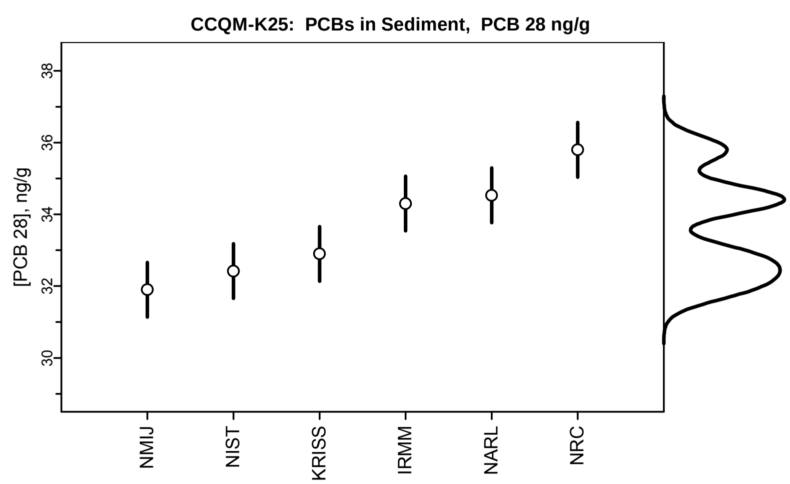


Figure 6: If all of the $U_{95}(x_i)$ were as small as the smallest reported value…

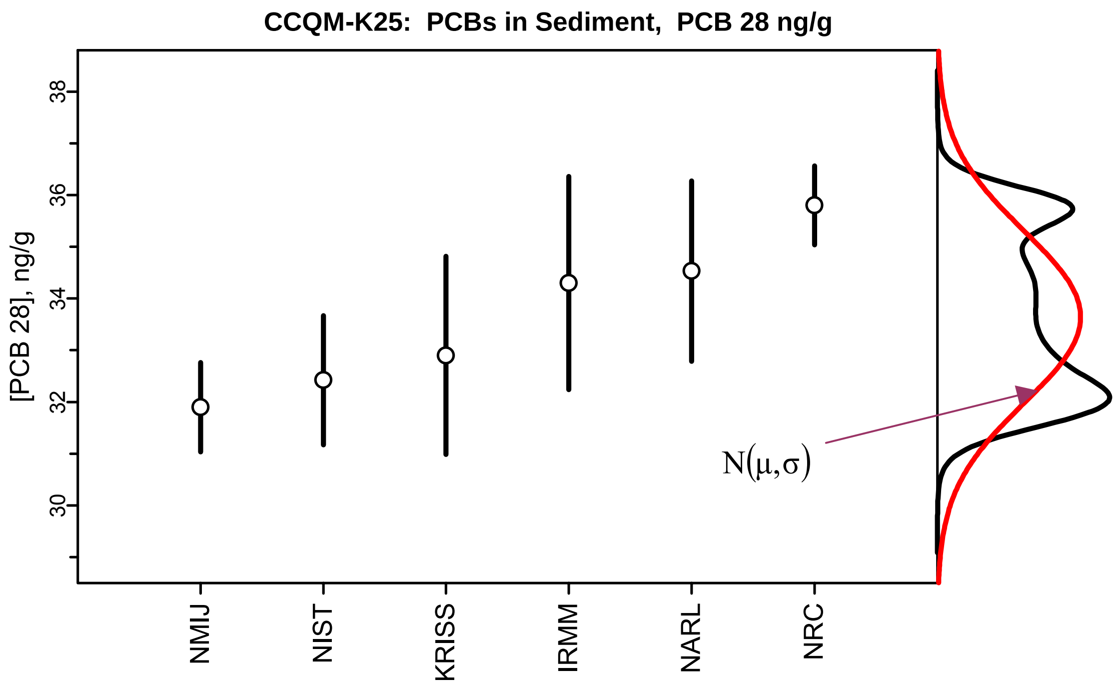


The use of MM-PDFs defined using individualized kernel dispersions, whether based upon precision estimates from replicate measurements or upon uncritically-evaluated estimates of uncertainty, has been strongly criticized [2]. As Figures 4 to 6 demonstrate, inferring metrological consequence from the shape of the MM-PDF strongly depends upon the quality of the estimated uncertainties. KCs appear to be the only current class of interlaboratory study that mandates reporting of measurement uncertainties of defined character: the $U_{95}(x_i)$. To the extent that a $N(x_i, U_{95}(x_i)/2)$ truly reflects the underlying states of knowledge of the location of the true value of the measurand, the MM-PDF truly summarizes KC results.

The MM-PDF is a useful tool for visually evaluating model distributions.

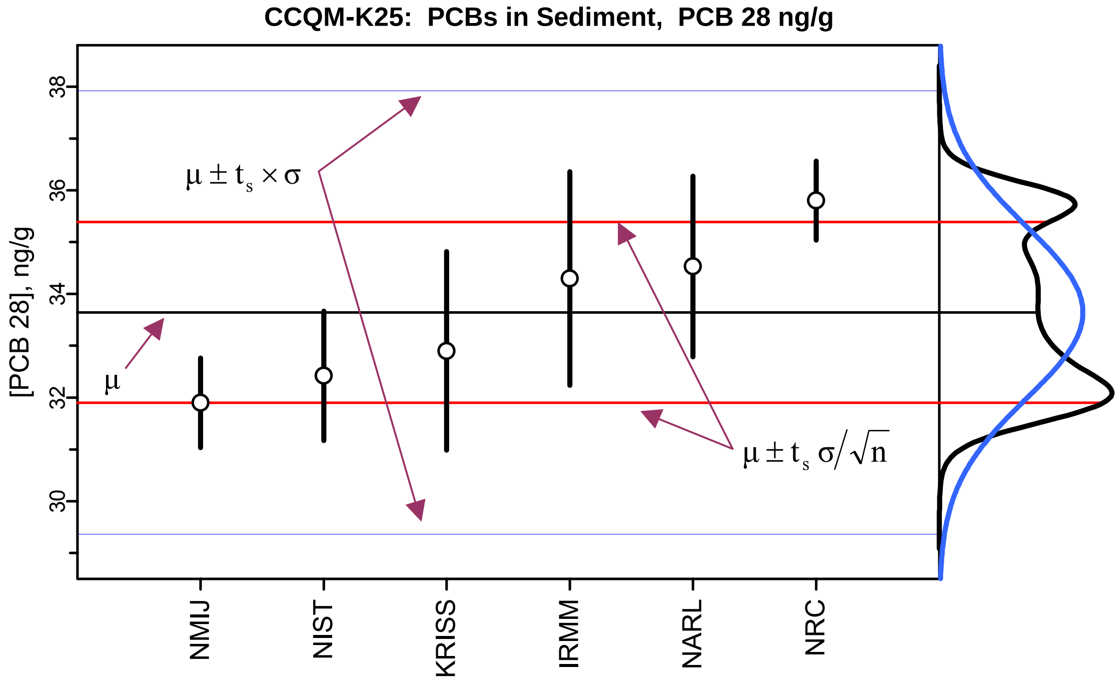
While the MM-PDF usefully summarizes a given set of KC measurements, it is difficult to use outside of a graphical context. Assuming that a set of results can plausibly be regarded as reflecting some underlying distribution of defined character, then it may be more efficient to summarize the results with the parameters required for the model distribution. Again, in the absence of other information, it is common practice to assume that results are normally distributed about a true value, μ, with a characteristic standard deviation, σ. Thus, the typical model is the normal kernel density N(μ,σ) (**Figure 7**).

Figure 7: Modeling KC Results as a N(μ,σ) Kernel Density



Once an appropriate model is established, decision lines characteristic of the model may provide visual aids for evaluating KC results, such as: μ, $\mu \pm t_s\sigma$ (the 95% confidence interval on the population of results where t_s is the appropriate Student's t statistic), and $\mu \pm t_s\, \sigma/\sqrt{n}$ (the 95% confidence interval on μ itself where n is the number of results) (**Figure 8**).

Figure 8: Model-Based Decision Lines: μ, $\mu \pm t_s\sigma$, $\mu \pm t_s\,\sigma/\sqrt{n}$



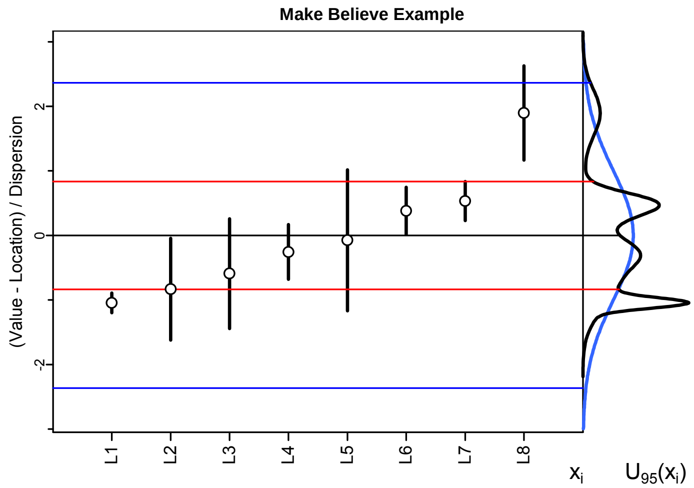
The inclusion of any part of an $x_i \pm U_{95}(x_i)$ interval within the $\mu \pm t_s\,\sigma/\sqrt{n}$ decision lines (as are all of those above) provides a visual evaluation of whether or not the particular result is in accord with the selected distributional model. Those results completely outside these lines will be metrologically interesting, however aggravating that finding would be to the participant(s) involved.

MM-PDF location and dispersion estimates utilize all of the information in a KC result.

Since the vast majority of interlaboratory results do not include meaningful estimates of uncertainty, few summary analysis techniques have been developed that can utilize all of the information provided in a set of KC $\{x_i, U_{95}(x_i)\}$ results. While not exhaustive, it is convenient to classify location and dispersion estimates into the following four groups:

- those that ignore the $U_{95}(x_i)$ (e.g., mean and standard deviation),
- those that use the $U_{95}(x_i)$ as weighting functions (e.g., uncertainty-weighted mean and standard deviation of the uncertainty-weighted mean),
- those that bootstrap-resample from $N(x_i, U_{95}(x_i)/2)$ kernels,
- those that characterize the MM-PDF built from the $N(x_i, U_{95}(x_i)/2)$ kernels.

The following "make believe" (CCQM KC data rescaled for graphical convenience) are used to illustrate some of the major properties of these representative statistics from these groups, using analogies to physical measurement methods where practical.



	x_i	$U_{95}(x_i)$
L1	2.020	0.161
L2	2.470	0.836
L3	2.984	0.900
L4	3.692	0.450
L5	4.077	1.157
L6	5.042	0.386
L7	5.364	0.322
L8	8.257	0.772

Location

Value		Parameter
4.24	◉	Mean
3.17	○	Uwt-mean
4.23	○	MPwt-mean
3.88	○	Median
3.27	○	Shorth
3.95	○	MM-median
2.98	○	MM-sh/mid
2.89	○	MM-sh/med
2.03	○	MM-mode
4.67	○	BS-mean
4.01	○	H15
4.07	○	A15

2.03 Minimum
4.67 Maximum

Dispersion

Value		Parameter
2.12	◉	S & ExU
1.85	○	Suwt & ExU
2.01	○	SMPwt & ExU
2.03	○	MADe & ExU
1.82	○	IQR & ExU
1.38	○	S(Shorth) & ExU
2.00	○	S(MM-median)
1.83	○	S(MM-shorth)
2.51	○	S(BS-mean)
2.28	○	S(BS-SD)
1.87	○	S(H15) & ExU

1.38 Minimum
2.51 Maximum

Expected Uncertainty

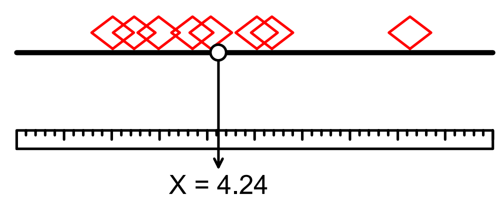
Value		Parameter
0.70	◉	Pooled Uc
0.61	○	Median Uc
0.00	○	Zero

Value	Parameter	
8	Number of {value, uncertainty} pairs	
0.70	Pooled Uc	*Expected Unc*
2.36	Student's t(1-0.95)	
4.24	Mean	*Location*
2.12	S & ExU	*Dispersion*
5.01	~U95(population): dispersion*ts	
1.77	~U95(location): dispersion*ts/sqrt(N)	
50	%RSD	

Estimates of Location

Estimates of location can be envisioned as identifying a particular value that, in some manner, balances two halves of the data. The following location estimates are intended to be representative of the more commonly used of the "ignore" and "weight" classes listed above, as well as presenting the bootstrap and MM-PDF methods evaluated for this study. Some of the definitions presented below require that none of the x_i have exactly the same value; this is a limitation of the descriptions but not of the practical algorithms.

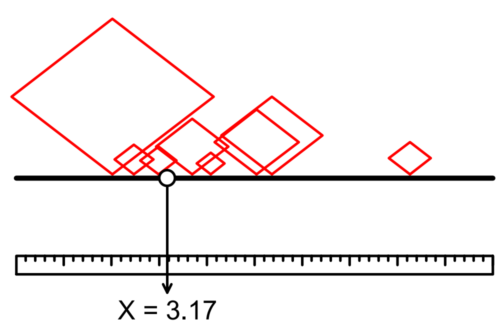
Mean: Find X such that $\sum_{i}(x_i - X) = 0$.



This is the by far the most widely used estimate of location. It can be regarded as the point of balance of a set of equally-weighted values on a beam. The $U_{95}(x_i)$ are not used in any way. Values have increasing influence on location as they deviate from the mean value. The mean is thus not a robust estimate; i.e., a small change in a single outlier value can greatly change the location of the balance point.

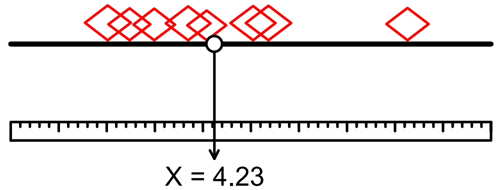
Uncertainty Weighted Mean, **Uwt-mean**:

Find X such that $\sum_{i}\frac{(x_i - X)}{\left(U_{95}(x_i)/2\right)^2} = 0$.



This is the most commonly used weighted location estimate. It can be regarded as the point of balance of a set of values that have weight equivalent to the squared-reciprocal of the $U_{95}(x_i)$. Although it has been advocated for use in KC evaluations [7], small errors in determination of the uncertainty may result in large changes in the estimated location. The Uwt-mean is thus robust to neither outlier x_i nor inappropriately small $U_{95}(x_i)$.

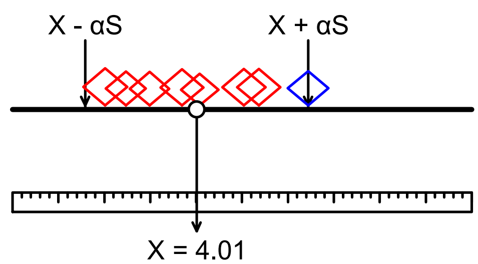
<u>Total Variance Weighted Means</u>, **MPwt-mean**:



Find X such that $\sum_i \frac{(x_i - X)}{\left(U_{95}(x_i)/2\right)^2 + s_{among}^2} = 0$.

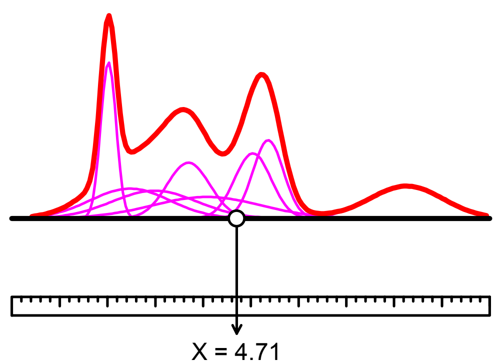
There are several variants of this estimate [8, 9]; only the original Mandel-Paule definition is included in this study. These methods iteratively estimate X and s_{among}, the among-participant standard deviation. For datasets such as that displayed here where s_{among} is relative large compared to the $U_{95}(x_i)$, the resultant location is close to the mean; when s_{among} is small (that is, the x_i are in excellent accord), the location will be close to the Uwt-mean. While protected from the influence of inappropriately small $U_{95}(x_i)$, the MPwt-mean and its relatives are not robust to outlier x_i.

<u>Winsorized Means</u>: Find X such that $\sum_i (z_i - X) = 0$



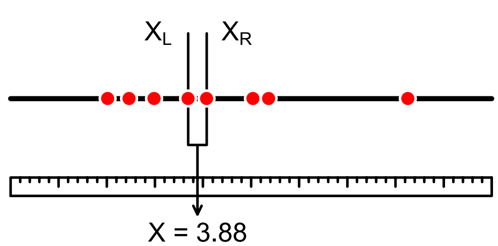
where $z_i = x_i$ if $|x_i\text{-}X| \le \alpha\, s_{among}$ and $z_i = X + \alpha\, s_{among}$ if $|x_i\text{-}X| > \alpha\, s_{among}$ and α is an empirical constant. The **A15** estimate uses an externally established estimate of s_{among}; the **H15** (Huber's Estimate 2) iteratively estimates X and s_{among}. While they do not use the $U_{95}(x_i)$ in any way, these methods are robust to outlier x_i and have been strongly recommended for use with interlaboratory results [10, 11]. These robust statistics are available in a (free!) spreadsheet add-in [12].

<u>Bootstrap Mean</u>, **BS-mean**: Find X such that



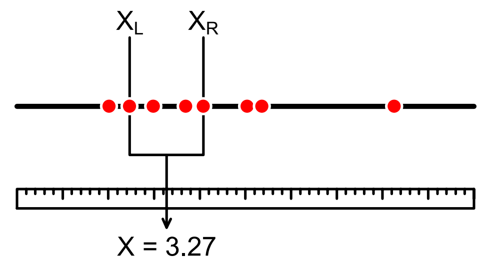
$\int_{-\infty}^{\infty} \frac{(x - X)}{\text{PDF}(x)^2} dx = 0$. In practice, this is evaluated as the mean of a large number of bootstrap-resampled x drawn from the MM-PDF [13]. While not infrequently used in the evaluation of multivariate models [14, 15], these bootstrap techniques do not appear to have been used in an interlaboratory context. The BS-mean fully utilizes the $U_{95}(x_i)$ and is robust to inappropriately small $U_{95}(x_i)$; however, it is somewhat sensitive to inappropriately *large* $U_{95}(x_i)$, particularly when associated with an outlier x_i.

<u>Median</u>: Find X_L and X_R, $X_L \le X_R$, such that

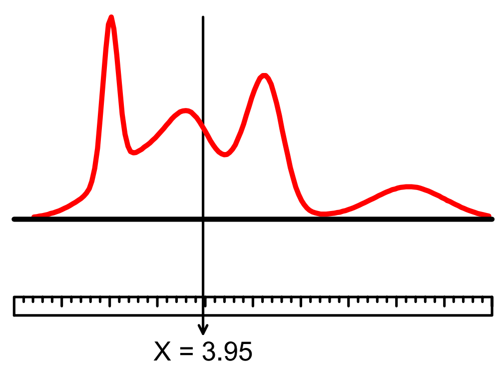


$\sum_{\forall x_i \le X_R} 1 = \sum_{\forall x_i \ge X_L} 1$. When $X_L \ne X_R$ (i.e., when there are an even number of data), the median is conventionally specified as $X = \frac{X_L + X_R}{2}$. The median is the most commonly used robust location estimate. It does not utilize $U_{95}(x_i)$, but it is robust to a moderate fraction of outlier x_i. The median has been proposed as a method for determining KCRVs [16, 17].

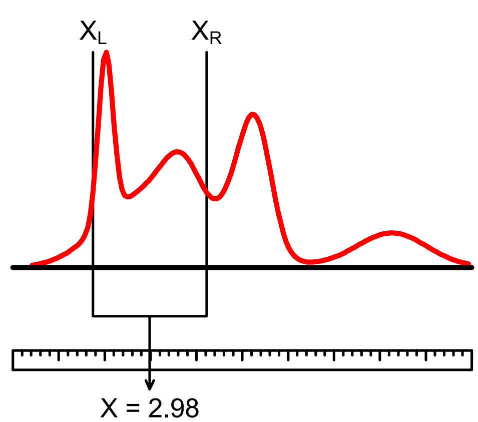
Shorth: Find X_L and X_R, $X_L \le X_R$, such that
$$\sum_{\forall X_L \le x_i \le X_R} 1 = \sum_{\forall x_i < X_L} 1 + \sum_{\forall x_i > X_R} 1$$ and $X_R - X_L$ is minimized. The shorth (shortest half) is $X = \dfrac{X_L + X_R}{2}$. This univariate analogue of the multivariate minimum volume ellipsoid estimator [18] is similar to the median in that it divides the number of data into two equal parts – but "interior" vs. "exterior" rather then "left" vs. "right". Conventions for treating odd numbers of data have been established. Like the median, it does not utilize $U_{95}(x_i)$ in any way and is robust to a moderate fraction of outlier x_i. Since the distance between X_R and X_L is much larger than in the case of the median, the location of the half-width X is less influenced by the detailed distribution of the interior or "inlier" x_i. When the separation between the ordered data, $x_{i+1} - x_i$, is the same for many of the inliers, there may exist multiple – equally valid – shorth estimates. This indeterminacy can, in principle, be avoided by combining all equivalent shorths. The major utility of the shorth is that it does not require that the data be symmetrically distributed about the estimated location and so may be more appropriate then the median when there are strong scientific grounds (e.g., asymmetrical biases such as extraction or degradation) for believing that the reported values can only asymptotically approach the true value of the measurand rather then being symmetrically distributed about the true value. The shorth has been used in the evaluation of interlaboratory studies involving high-level metrology [19].



Mixture Model Median, **MM-median**: Find X such that $\int_{-\infty}^{X} \mathrm{PDF}(x)\mathrm{d}x = \int_{X}^{\infty} \mathrm{PDF}(x)\mathrm{d}x$. This estimate is closely related to the median, but proceeds by dividing the MM-PDF by area rather than the x_i by number. Like the median, it is robust to outlier x_i; unlike the median, it uses information provided in the $U_{95}(x_i)$. This estimate appears to be relatively novel to both the metrological and statistical literature.

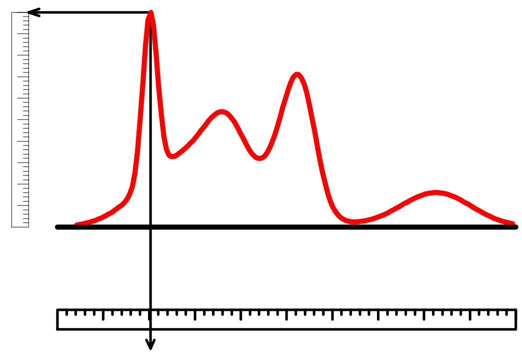


Mixture Model Shorth, **MM-shorth**: Find X_L and X_R, $X_L \le X_R$, such that
$$\int_{X_L}^{X_R} \mathrm{PDF}(x)\mathrm{d}x = \int_{-\infty}^{X_L} \mathrm{PDF}(x)\mathrm{d}x + \int_{X_R}^{\infty} \mathrm{PDF}(x)\mathrm{d}x$$ and that minimize $X_R - X_L$. Like the MM-median, this is closely related to the shorth. It is robust to outlier x_i and uses information provided in the $U_{95}(x_i)$. It shares the major strength of the shorth in that it does not require the exterior density to be



symmetrically distributed. It also shares a tendency to have multiple equivalent values when the MM-PDF has two or more modes of equal area. Two variants for combining X_L and X_R into a single location estimate have been evaluated. In direct analogy to the shorth, **MM-sh/mid** is defined as $X = \frac{X_L + X_R}{2}$; to more fully exploit the detailed shape of the interior density of the MM-PDF, the **MM-sh/med** is the X such that $\int_{X_L}^{X} \mathrm{PDF}(x)\mathrm{d}x = \int_{X}^{X_R} \mathrm{PDF}(x)\mathrm{d}x$. Like the MM-median, both of these estimates appear to be relatively novel.

Mixture Model Mode, **MM-mode**: Find X where MM-PDF(x) is largest. While the mode has been advocated for use in the evaluation of interlaboratory data [20] and the MM-mode uses information provided in the $U_{95}(x_i)$, as with the Uwt-mean this estimate is sensitive to small outlier $U_{95}(x_i)$. If it has a strength beyond its simplicity, it is that – like the shorth and MM-shorth – it does not require that the density be symmetrically distributed about the dividing location.



Relationships Among the Location Estimates

Table 1: Some Characteristics of Location Estimates

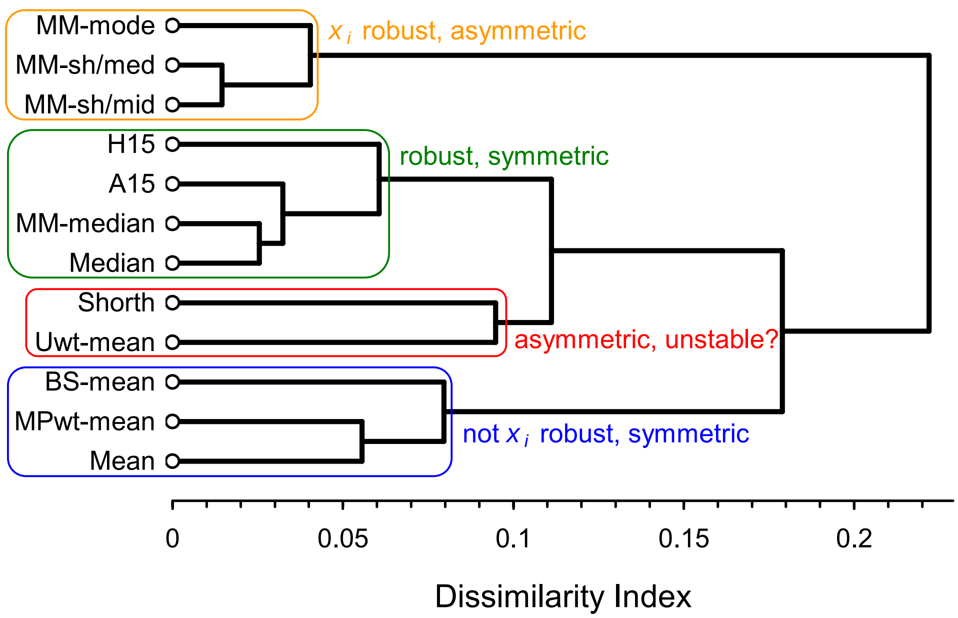
Estimate	Literature Characterized	Uses $U_{95}(x_i)$	Robust to outlier x_i	Robust to outlier $U_{95}(x_i)$	Assumes Symmetry
Mean	Yes	No	No		Yes
Uwt-mean	Yes	Yes	No	No	No
MPwt-mean	Yes	Yes	No	Yes	No
A15	Yes	No	Yes		Yes
H15	Yes	No	Yes		Yes
BS-mean	No	Yes	No	Yes	Yes
Median	Yes	No	Yes		Yes
Shorth	Yes	No	Yes		No
MM-median	No	Yes	Yes	Yes	Yes
MM-sh/mid	No	Yes	Yes	Yes	No
MM-sh/med	No	Yes	Yes	Yes	No
MM-mode	No	Yes	Yes	No	No

Table 1 summarizes some characteristics of interest to the determination of KCRV for the various location estimates discussed above. However, neither these properties nor the method of computation indicate the kind or extent of functional relationships, if any, among the summary location values that they yield. The existence of such relationships is suggested from cluster analysis [21] of normalized differences from the KCRV, $z_i = (x_i – \mathrm{KCRV})/\mathrm{KCRV}$, for some published KC data (**Figure 9**).

There appear to be three groups of estimates that are closely related:

- mean, MPwt-mean, and BS-mean. These estimates are not robust to x_i outliers and assume that the data are symmetrically distributed about a balance point.
- median, MM-median, A15, and H15. These estimates are robust to x_i outliers and assume that the data are symmetrically distributed about a balance point.
- MM-sh/mid, -sh/med, and –mode. These estimates are robust to x_i outliers and do not assume that the data are symmetrically distributed about a balance point.

Figure 9: Relationships Among Location Estimates
for 21 Organic Chemical Measurands Measured in CCQM KCs.



(The Dissimilarity Index is $\sqrt{2\left(1-R_{jk}^2\right)}$ where R_{jk} is Pearson's correlation between variables j and k. The larger the Dissimilarity Index, the less similar the behavior of the variables.)

The remaining, fairly weak, grouping of the shorth and the Uwt-mean is not as easily interpreted, although neither estimate assumes that the (un-weighted) data are symmetrically distributed about a balance point. It is interesting that the shorth and the MM-shorth estimates are not at all closely related, unlike the median and MM-median. It may be that the shorth and Uwt-mean provide relatively unstable location estimates for these characteristically few number-of-data studies.

Modeling Measurement Dispersion: Among, Within, and Total

There are two sources of dispersion in KC data: that resulting from differences among the x_i, s_{among}, and that reflecting the participants' measurement uncertainty, s_{within}. Estimates based on $N(x_i, U_{95}(x_i)/2)$ kernels "see" the total dispersion, $s_{\text{total}} = \sqrt{s_{\text{among}}^2 + s_{\text{within}}^2}$. Estimates that ignore the $U_{95}(x_i)$ or that use them as weights "see" only s_{among}; these need to be augmented with an independent estimate of s_{within} before they can be compared with the others. The usual estimate for s_{within} is the "pooled" (average of the sum-of-the-squared) uncertainty estimates, $\sqrt{\frac{1}{n}\sum_i^2 (U_{95}(x_i)/2)^2}$. Other estimates, such as the median $U_{95}(x_i)/2$, may be appropriate when there are very large outlier $U_{95}(x_i)$.

Because standard deviations add in quadrature, only when $s_{\text{within}} / s_{\text{among}} > 0.5$ (that is, the spread among the x_i is no more than twice the average spread of the stated uncertainties) will s_{total} be much (10%) larger than s_{among}. **Figure 10** displays highly concordant KC results modeled as N(μ, s_{among}); **Figure 11** displays the same data modeled as N(μ, s_{total}). Comparison with the empirical MM-PDF suggests that the N(μ, s_{among}) model indeed underestimates the total dispersion. (The Figures also suggest that most participants here either overestimated their $U_{95}(x_i)$ or use procedures that have very similar biases.)

Figure 10: Highly Concordant KC Results Modeled as a N(μ, s_{among}) Distribution

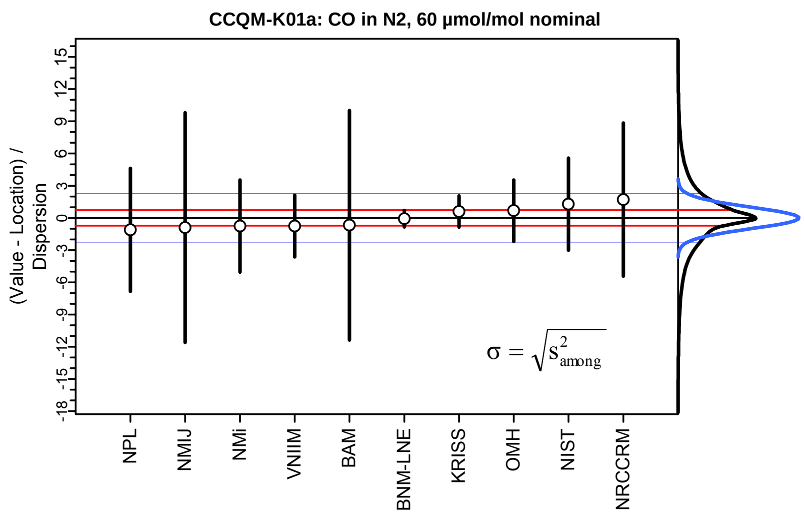
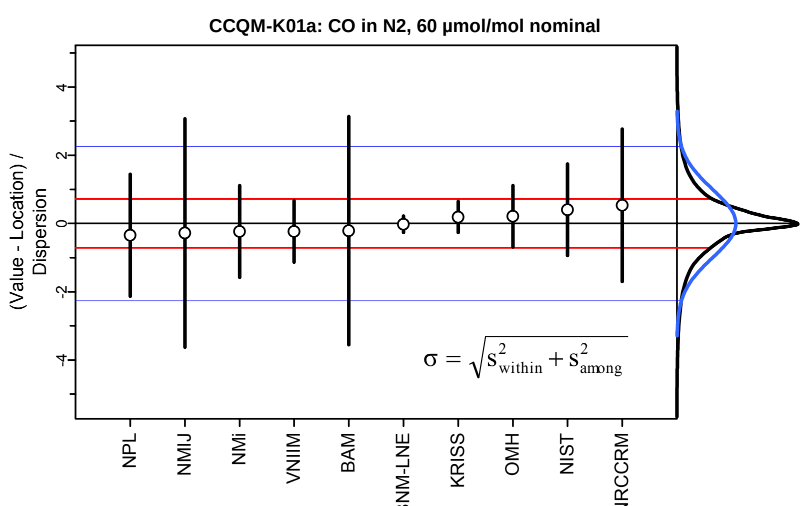


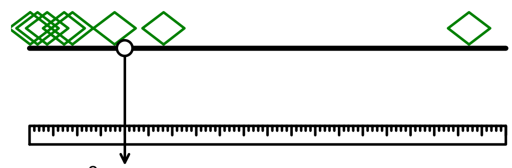
Figure 11: Highly Concordant KC Results Modeled as a N(μ, s_{total}) Distribution



Estimates of Dispersion

Dispersion estimates are less cogently envisioned in analogy to physical measurements as they address the spread of the data about a specified location rather then a set of fixed locations. Some of the estimates can be represented as a balance point of square deviations,; there is no fixed rank order of the original data nor fixed highest value. The two "beam balance" cartoons below display the locations of the $(x_i - \mathrm{X})^2$, using the same "make believe data as used for the location estimates, as green diamonds; the statistic of interest is the square-root of the balance point. Other estimates can be represented as either a "pan balance" division similar to that presented for the median or as length measurements; in both cases, the statistic of interest must be scaled to give an appropriate estimate of dispersion for the N(μ,σ) model.

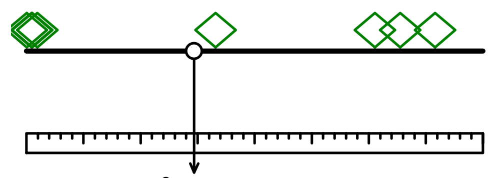
Standard Deviation, **S**: $\sqrt{\frac{1}{n-1}\sum_i^n (x_i - \mathrm{X})^2}$.



The standard deviation (here specified as the standard deviation of a sample since one degree-of-freedom is consumed in estimating the mean) is by far the most common dispersion estimate. The $U_{95}(x_i)$ are not used in any way. Large squared deviations have great influence, making S very much less robust than the mean. The estimate assumes that the x_i are symmetrically distributed about the mean.

Uncertainty Weighted Standard Deviation, **Suwt**:

$$\sqrt{\frac{1}{(n-1)}\left(\sum_i^n (x_i - \mathrm{X})^2 \Big/ \frac{(U_{95}(x_i)/2)^2}{n\sum_i^n (U_{95}(x_i)/2)^2}\right)} .$$



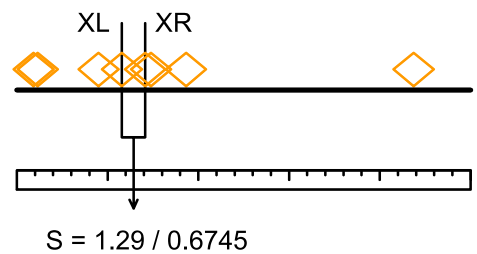
This statistic is typically computed in the form of a standard deviation of weighted mean, a result of the conventional scaling of the weights such that their sum is 1.0. The standard deviation of the population results is recovered by scaling the weights such that their sum is 1/n. While the $U_{95}(x_i)$ are used, they only influence the s_{among} balance point and do not add additional dispersion. Like the Uwt-mean, Suwt is robust to neither outlier x_i nor inappropriately small $U_{95}(x_i)$. Because of the variable weighting, the x_i are not necessarily assumed to be symmetrically distributed about the mean.

Total Variance Weighted Standard Deviations, **SMPwt**: As with the related location statistics, there are several variants of this estimate [8, 9]. Only the original Mandel-Paule definition is included in this study, modified to yield a standard deviation rather than a standard deviation of the MPwt-mean. There is no closed-form formula; the location and dispersion statistics are calculated by iteration. For datasets where s_{among} is large compared to s_{within} SMPwt will be similar to S; when s_{among} is small compared to s_{within} SMPwt will be similar to Suwt. In neither case does the use of the $U_{95}(x_i)$ as part of the weighting function add addition dispersion. While protected from the influence of inappropriately small $U_{95}(x_i)$, the SMPwt and its relatives are not robust to outlier x_i.

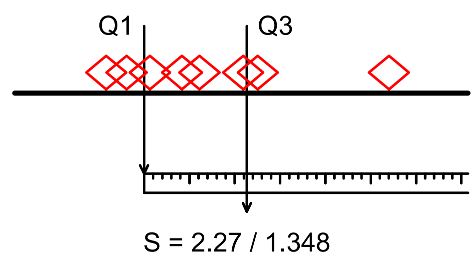
Winsorized Standard Deviation, **S(H15)**: This statistic is iteratively calculated in conjunction with the **H15** location estimate. It is robust to outlier x_i. In the absence of outlier x_i it is identical to S; it does not use the $U_{95}(x_i)$. The estimate assumes that the x_i are symmetrically distributed about the mean. It has been strongly recommended for use with interlaboratory results [11]. This statistic is available as an spreadsheet add-in [12].

Bootstrap Standard Deviations. Two variant statistics that are readily calculated from the bootstrap resampling data used for the BS-mean have been studied. **S(BS-SD)** is $\sqrt{\sum_m^M S_m^2 \Big/ M}$, the pooled standard deviation of a large number, M, of standard deviations of bootstrap-resampled data drawn from the MM-PDF distribution, S_m. **S(BS-mean)** is $\sqrt{M\sum_m^M \left(X_m - \mathrm{X}\right)^2 \Big/ \left(M-1\right)}$, the standard deviation of a large number of BS-mean estimates, X_m, expanded to estimate the standard deviation of the population rather than the standard deviation of the BS-mean. Both estimate s_{total}. As with the BS-mean, neither is robust to outlier x_i nor large $U_{95}(x_i)$. Both assume that the data are symmetrically distributed.

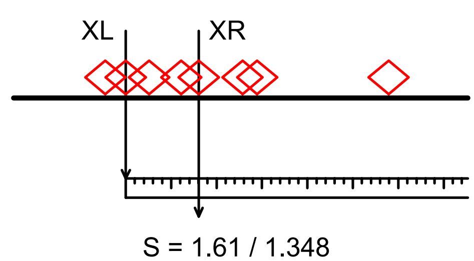
Median Absolute Deviation About the Median, **MADe**. The MADe is calculated in the same manner as the median but uses the absolute deviations from the median rather than the x_i, MEDIAN($|x_i$-X$|$) / 0.6745. The constant 0.6745 is based upon relationships between absolute deviations and the standard deviation for normal distributions [16]. The MADe thus uses the median of the $|x_i$-X$|$ to estimate the distribution the x_i would have if the x_i were all true members of single normal distribution. The MADe does not use the $U_{95}(x_i)$ and only estimates s_{among}. It assumes that the x_i are symmetrically distributed about the median.



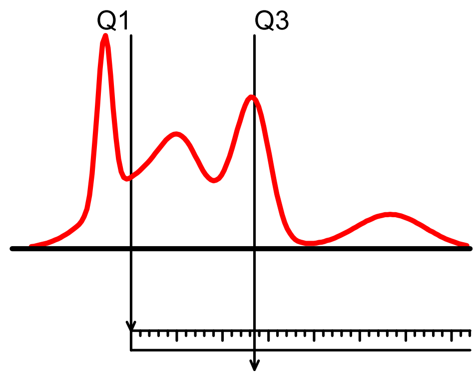
Interquartile Range-based Standard Deviation, **IQR**. The IQR is calculated from the span of the central 50% of the x_i, $\{Q_3(x_i)$-$Q_1(x_i)\}$ / 1.348, where Q_1 is the location of the first quartile and Q_3 is the location of the third quartile of the x_i. The constant 1.348 is based upon the relationship between the spans about the center of a normal distribution that include 68% and 50% of the distribution [22]. The IQR thus uses the span of the central 50% of the data to estimate the distribution the x_i would have if the x_i were all true members of single normal distribution. The IQR does not use the $U_{95}(x_i)$ and only estimates s_{among}. It assumes that the x_i are symmetrically distributed about the median.



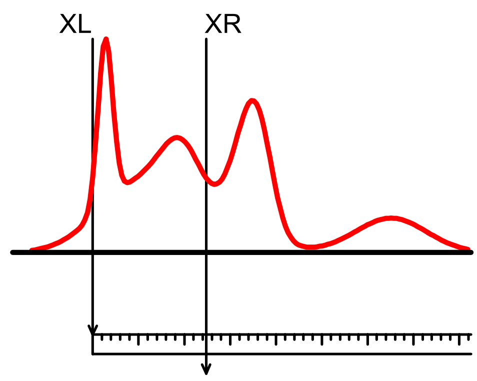
Shorth-based Standard Deviation, **S(shorth)**. The S(Shorth) is calculated in the same manner as the IQR, replacing the span of the interquartile range with the span of the shortest half: $(X_R$-$X_L)$ / 1.348 [22]. The S(shorth) uses the span of the most compact 50% of the data to estimates the distribution the x_i would have if the x_i were all true members of single normal distribution. The



S(shorth) does not use the $U_{95}(x_i)$ and only estimates s_{among}. It does not assume that the x_i are symmetrically distributed about a central location.



MM-median-based Standard Deviation, **S(MM-median)**.
The S(MM-median) is calculated in the same manner as the IQR, replacing the span of the interquartile range with the span of the central 50% of the MM-PDF density: $(Q_3(\text{MM-PDF}(x_i)) - Q_1(\text{MM-PDF}(x_i)) / 1.348$ [22]. The S(MM-median) thus estimates the distribution the x_i would have if the x_i were all true members of single normal distribution. The S(MM-median) uses the $U_{95}(x_i)$. It estimates s_{total} and assumes that the x_i are symmetrically distributed about a central location.



MM-shorth-based Standard Deviation, **S(MM-shorth)**. The S(MM-shorth) is calculated in the same manner as the S(shorth), replacing the span of the shortest-half with the span of the most compact 50% of the MM-PDF density: $(X_R - X_L) / 1.348$ [22].. The S(MM-shorth) thus estimates the distribution the x_i would have if the x_i were all true members of single normal distribution. The S(MM-shorth) uses the $U_{95}(x_i)$. It estimates s_{total} and does not assume that the x_i are symmetrically distributed about a central location.

Relationships Among the Dispersion Estimates

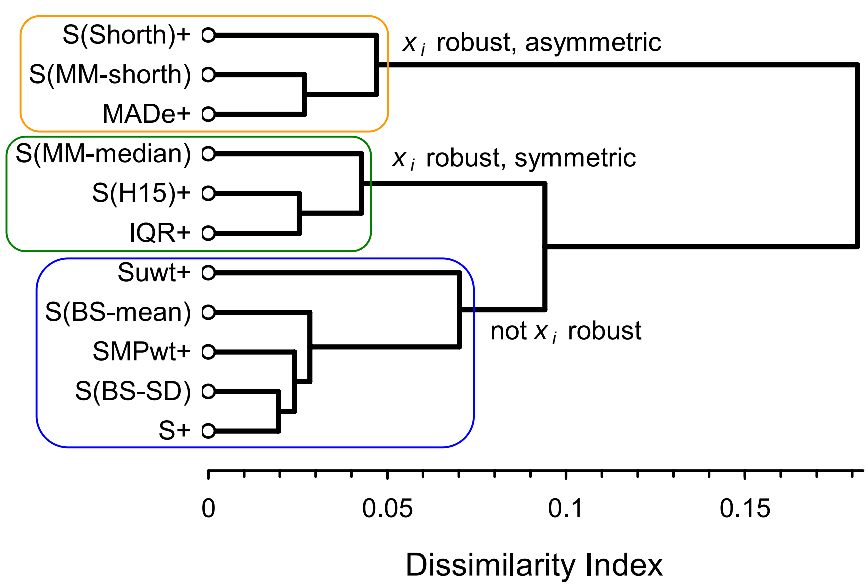
Table 2: Some Characteristics of Dispersion Estimates

Estimate	Literature Characterized	Uses $U_{95}(x_i)$	Robust to outlier x_i	Robust to outlier $U_{95}(x_i)$	Assumes Symmetry
S	Yes	No	No		Yes
Suwt	Yes	Yes	No	No	No
SMPwt	Yes	Yes	No	Yes	No
S(H15)	Yes	No	Yes		Yes
S(BS-mean)	No	Yes	No	No	Yes
S(BS-SD)	No	Yes	No	Yes	Yes
MADe	Yes	No	Yes		No
IQR	Yes	No	Yes		No
S(shorth)	Yes?	Yes	Yes	Yes	No
S(MM-median)	No	Yes	Yes	Yes	Yes
S(MM-shorth)	No	Yes	Yes	Yes	No

Table 2 summarizes some characteristics of interest to the determination of u(KCRV) with the various dispersion estimates discussed above. As with the location estimates, neither these properties nor the method of computation indicate the kind or extent of functional relationships, if any, among the summary dispersion values that they yield. The existence of such relationships is suggested from cluster analysis [21] of normalized differences from the u(KCRV), z_i = (S-u(KCRV))/KCRV, for some published KC data (**Figure 12**). Once again, there appear to be three groups of estimates that are closely related (statistics augmented with s_{within} are denoted with "+" concatenated to the code name):

- S+, S(BS-SD), SMPwt+, S(BS-mean), and (perhaps), Suwt+. These estimates are not robust to x_i outliers. The close similarity of the augmented S to the bootstrap S(BS-SD) suggests that the augmentation process is appropriate.
- IQR+, S(H15)+, and S(MM-median). These estimates are robust to x_i outliers and assume that the data are symmetrically distributed about a balance point.
- MADe+, S(MM-shorth), and S(shorth)+. These estimates are robust to x_i outliers and do not assume that the data are symmetrically distributed about a balance point.

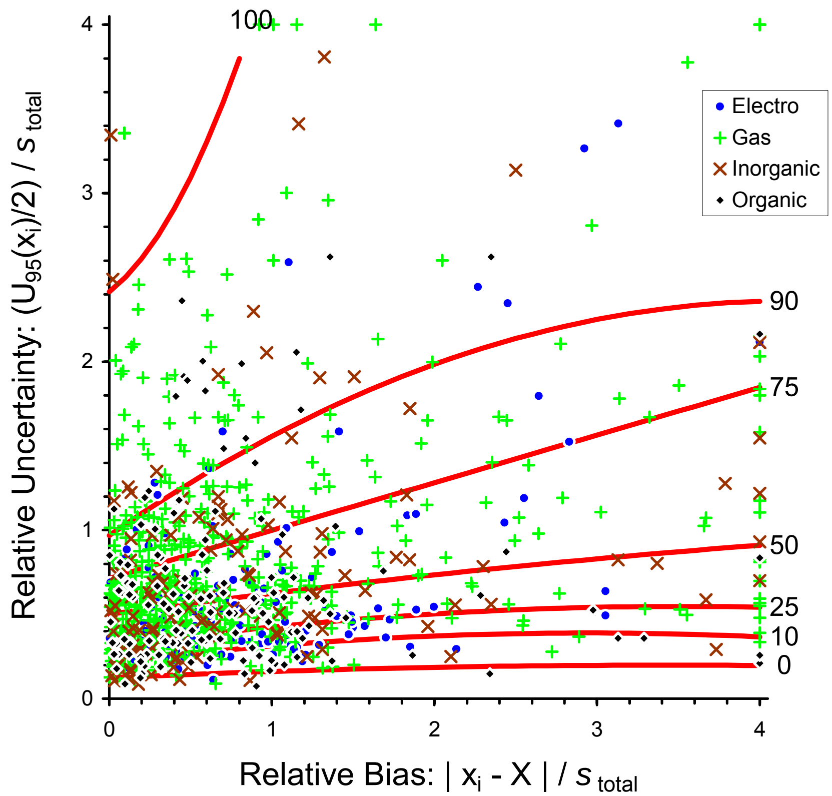
Figure 12: Relationships Among Dispersion Estimates
for 21 Organic Chemical Measurands Measured in CCQM KCs.



Weighted Statistics Make an Unsupported Assumption

All statistics that use the $U_{95}(x_i)$ as part of a weighting function presuppose the existence of a strong positive relationship between the magnitude of the uncertainty, $U_{95}(x_i)$, and the magnitude of the bias of the x_i from the true value, $|x_i\text{-X}|$. If there is no such relationship, at least for the measurements having the smaller $U_{95}(x_i)$, such weighting has little or no practical utility. **Figure 13** empirically explores this for the current CCQM KC results by plotting the relative uncertainties, $(U_{95}(x_i)/2)/s_{\text{total}}$, as a function of the associated relative biases, $|x_i\text{-X}|/s_{\text{total}}$. The MM-median and S(MM-shorth) have been used to estimate the X and s_{total} parameters.

Figure 13: Relative Uncertainties of CCQM KC Results as a Function of Relative Bias.



The red lines approximate the percentage of the results that are between the line and the origin as a function of the relative bias. If there were no relationship between the relative uncertainties and biases, all lines would be horizontal; if there were a strongly positive relationship, the 25%, 50%, and 75% lines would be expected to have significant positive slope. Any functional relationship that exists in these data is not particularly strong.

What Are the Best Statistics for Evaluating KC Results? It depends!

Before the most appropriate statistics for establishing KCRV and u(KCRV) can be identified, the intended use for the summary information must be clearly identified. There are at least three potential uses of KC information:

- Inference of the true value of the measurand in a material. While more appropriate for pilot studies and during the validation stages of KC studies, there has been forceful arguments on the need for a KCRV to represent the measurand true value [5]. If a true value is to be inferred from KC results, summary statistics that are robust to outliers and include all known sources of dispersion are required. If there are no metrological grounds for assuming asymmetric distribution of bias, the {median, MADe+} and {MM-median, S(MM-shorth)} appear to provide similar results. If the biases are known to be strongly asymmetrical, the {MM-shorth, S(MM-shorth)} are more reliable then the {shorth, S(shorth)} with small data sets.
- Inference of the performance characteristics of metrology used for a given measurand as practiced by NMIs for establishing appropriate traceability. While not an explicit

goal, KC data implicitly provide guidance on the limits of current metrological practice for measurands of considerable importance. Characterizing "best case" metrology also requires summary statistics that are robust to outliers but may not require inclusion of the s_{within} dispersion component. Any of the robust location statistics are suitable. The MADe appears most appropriate if the s_{among} dispersion is of primary interest; the MADe+ and S(MM-shorth) are appropriate when s_{total} dispersion is of primary interest.

- Description of the performance of all official participants in a KC. To truly describe all results, no data should be excluded nor differentially weighted and all sources of dispersion should be included. The {BS-mean, S(BS-SD)} may provide the most complete evaluation of KC results, but they are not yet well characterized nor established in the literature. The {mean, S+} are appropriate and have the great advantage of being well established.

Example analyses for some combinations of location and dispersion estimates are located at the end of the References.

Acknowledgement

I thank Reenie M. Parris and John C. Travis of NIST for their encouragement, questions, and rapid editing; Jim J. Filliben of NIST for taking some of my ideas seriously; David Banks of Duke University for excellent ideas and his tolerance of silly questions; and Steve L.R. Ellison of LGC for making robust methods suitable for use with interlaboratory comparison data easily and freely accessible. In particular I thank (I think) Wille E. May for getting me involved in the analysis of CCQM results and for his support of my investigations.

References

1 B.W. Silverman, *Density Estimation for Statistics and Data Analysis*, Chapman and Hall, London, 1986.

2 P.J. Lowthian, M. Thompson, Bump-hunting for the proficiency tester – searching for multimodality. Analyst 2002;127:1359-1364.

3 W.P. Cofino, I.H.M. van Stokkum, D.E. Wells, F. Ariese, J.-W. M. Wegener, R.A.L. Peerboom. A new model for the inference of population characteristics from experimental data using uncertainties. Application to interlaboratory studies. Chemometr. Intell. Lab. 2000;53(1-2):37-55.

4 J. de Boer, W.P. Cofino. First world-wide interlaboratory study on polybrominated diphenylethers (PBDEs). Chemosphere 2002;46:625-633.

5 Steele, A.G., Hill, K.D., Douglas, R.J. Data pooling and key comparison reverence values. Metrologia 2002, 39, 269-277.

6 P. Ciarlini, M. Cox, F. Pavese. G. Regoliosi. The use of a mixture of probability distributions in temperature interlaboratory comparisons. Metrologia 2004;41:116-121.

7 Hasselbarth, W., Bremser, W., Pradel, R. Uncertainty-based evaluation of certification study data. Fresen. J. Anal. Chem. 1998, 360, 317-21.

8 Paule, R.C., Mandel, J. Consensus values and weighting factors. J. Res. NBS 1982, 87(5), 377-85.

9 Rukhin, A.L., Biggerstaff, B.J., Vangel, M.G. Restricted maximum likelihood estimation of a common mean and the Mandel-Paule algorithm. J. Stat. Plan. Infer. 2000, 83(2), 319-30.

10 Analytical Methods Committee. Robust statistics - How not to reject outliers. Part 1. Basics. Analyst 1989, 114, 1693-8.

11 Analytical Methods Committee. Robust statistics - How not to reject outliers. Part 2. Inter-laboratory trials. Analyst 1989, 114, 1699-1702.

12 Ellison, S. RobStat.xla http://www.rsc.org/lap/rsccom/amc/amc_index.htm

13 Diaconis, P., Efron, B. Computer-intensive methods in statistics. Scientific American 1983, 248(5), 116-130.

14 Duewer, D.L., Kowalski, B.R., Fasching, J.L. Improving the reliability of factor analysis of chemical data by utilizing the measured analytical uncertainty. Analytical Chemistry 1976, 48, 2002-2010.

15 Faber, N.M. Uncertainty estimation for multivariate regression coefficients. Chemometr. Intell. Lab. 2002, 64(2), 169-179.

16 Müller, J.W. Possible advantages of a robust evaluation of comparisons. J. Res. NIST 2000;105(4):551-555.

17 Cox, M.G. The evaluation of key comparison data. Metrologia 2002, 39(6), 589-595.

18 Rousseeuw, P.J. Multivariate estimation with high breakdown point. In: Grossman, W., Pflug, G., Nincze, I., Wetrz, W. (eds.), *Mathematical Statistics and Applications*, 1985, 283-297. Reidel, Dordrecht, The Netherlands.

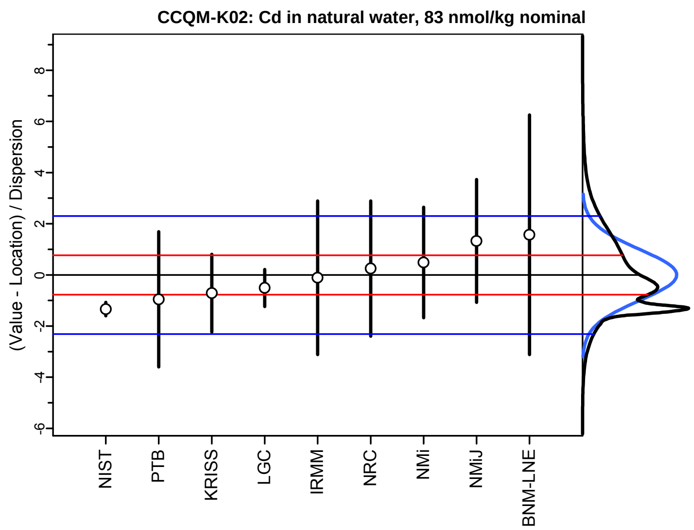
19 Rose, A.H., Wang, C.-M., Byer, S.D. Round Robin for optical fiber Bragg grating metrology. J. Res. NIST 2000, 105(6), 839-866.

20 Bickel, D.R. Robust estimators of the mode and skewness of continuous data. Comput. Stat. Data An. 2002, 39(2), 153-163.

21 Fahmy, T. Ascendant hierarchical cluster analysis, XLSTAT. http://www.xlstat.com.

22 Stuart, A., Ord, J.K. *Kendall's Advanced Theory of Statistics. Volume 1. Distribution Theory.* Sixth Edition, Chapter 10. Edward Arnold, London, 1994.

{Mean, S}



Location

Value		Parameter
83.5	●	Mean
82.5	○	Uwt-mean
82.9	○	MPwt-mean
83.4	○	Median
83.2	○	Shorth
83.3	○	MM-median
82.9	○	MM-sh/mid
82.9	○	MM-sh/med
82.4	○	MM-mode
83.6	○	BS-mean
83.5	○	H15
83.5	○	A15

82.4 Minimum
83.6 Maximum

83.5 KCRV
0.8 KCRU

Dispersion

Value		Parameter
0.8	●	S & ExU
0.4	○	Suwt & ExU
0.3	○	SMPwt & ExU
0.7	○	MADe & ExU
0.7	○	IQR & ExU
0.7	○	S(Shorth) & ExU
1.3	○	S(MM-median)
1.0	○	S(MM-shorth)
1.4	○	S(BS-mean)
1.4	○	S(BS-SD)
0.9	○	S(H15) & ExU

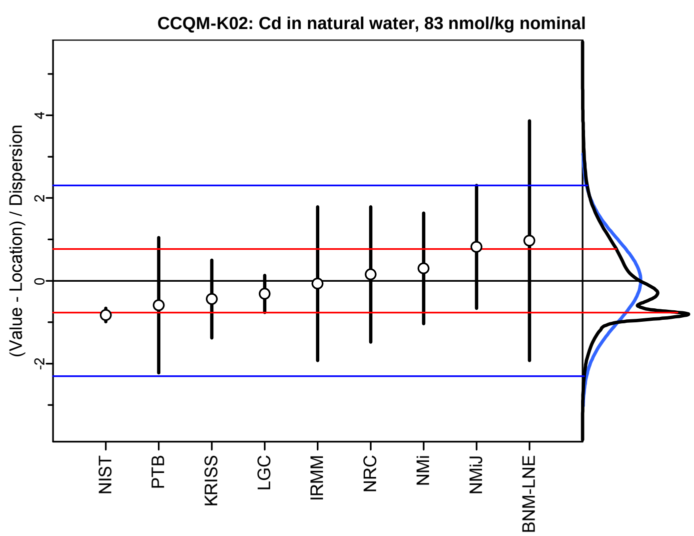
0.3 Minimum
1.4 Maximum

Expected Uncertainty

Value		Parameter
1.1	○	Pooled Uc
1.0	○	Median Uc
0.0	●	Zero

Value	Parameter	
9	Number of {value, uncertainty} pairs	
0.0	Zero	*Expected Unc*
2.3	Student's t(1-0.95)	
83.5	Mean	*Location*
0.8	S & ExU	*Dispersion*
1.9	~U95(population): dispersion*ts	
0.6	~U95(location): dispersion*ts/sqrt(N)	
1.0	%RSD	

{Mean, S+}



Location

Value		Parameter
83.5	●	Mean
82.5	○	Uwt-mean
82.9	○	MPwt-mean
83.4	○	Median
83.2	○	Shorth
83.3	○	MM-median
82.9	○	MM-sh/mid
82.9	○	MM-sh/med
82.4	○	MM-mode
83.6	○	BS-mean
83.5	○	H15
83.5	○	A15

82.4 Minimum
83.6 Maximum

83.5 KCRV
0.8 KCRU

Dispersion

Value		Parameter
1.3	●	S & ExU
1.1	○	Suwt & ExU
1.1	○	SMPwt & ExU
1.3	○	MADe & ExU
1.3	○	IQR & ExU
1.3	○	S(Shorth) & ExU
1.3	○	S(MM-median)
1.0	○	S(MM-shorth)
1.6	○	S(BS-mean)
1.5	○	S(BS-SD)
1.4	○	S(H15) & ExU

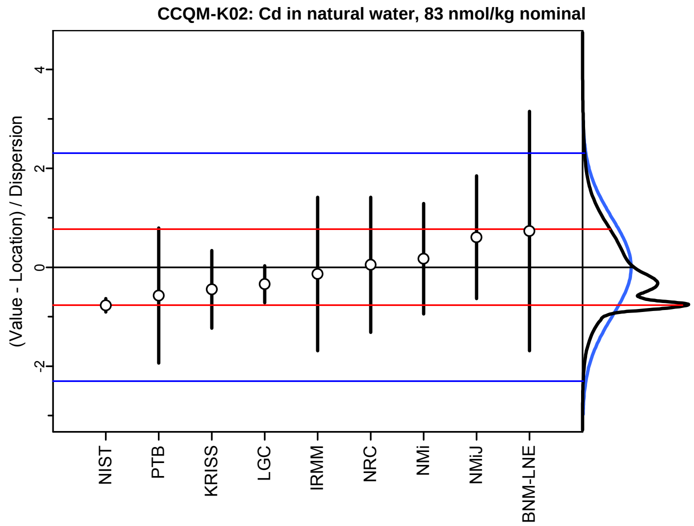
1.0 Minimum
1.6 Maximum

Expected Uncertainty

Value		Parameter
1.1	●	Pooled Uc
1.0	○	Median Uc
0.0	○	Zero

Value	Parameter	
9	Number of {value, uncertainty} pairs	
1.1	Pooled Uc	*Expected Unc*
2.3	Student's t(1-0.95)	
83.5	Mean	*Location*
1.3	S & ExU	*Dispersion*
3.1	~U95(population): dispersion*ts	
1.0	~U95(location): dispersion*ts/sqrt(N)	
1.6	%RSD	

{BS-mean, S(BS-mean)}



Location

Value	Parameter
83.5 ○	Mean
82.5 ○	Uwt-mean
82.9 ○	MPwt-mean
83.4 ○	Median
83.2 ○	Shorth
83.3 ○	MM-median
82.9 ○	MM-sh/mid
82.9 ○	MM-sh/med
82.4 ○	MM-mode
83.6 ◉	BS-mean
83.5 ○	H15
83.5 ○	A15

82.4 Minimum
83.6 Maximum

83.5 KCRV
0.8 KCRU

Dispersion

Value	Parameter
0.8 ○	S & ExU
0.4 ○	Suwt & ExU
0.3 ○	SMPwt & ExU
0.7 ○	MADe & ExU
0.7 ○	IQR & ExU
0.7 ○	S(Shorth) & ExU
1.3 ○	S(MM-median)
1.0 ○	S(MM-shorth)
1.6 ◉	S(BS-mean)
1.5 ○	S(BS-SD)
0.9 ○	S(H15) & ExU

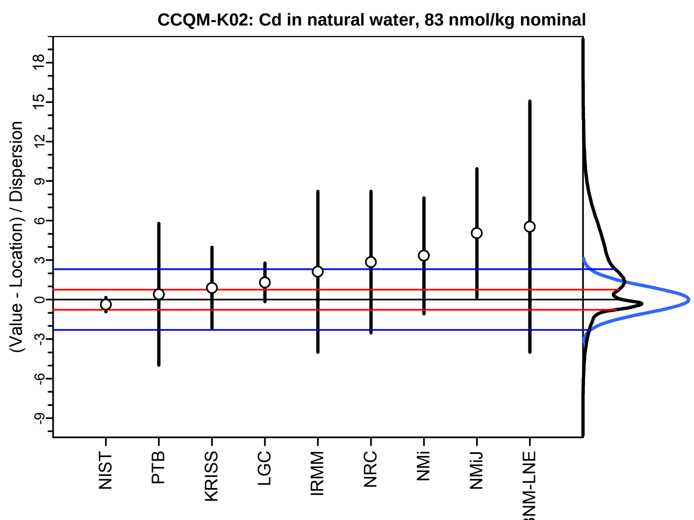
0.3 Minimum
1.6 Maximum

Expected Uncertainty

Value	Parameter
1.1 ○	Pooled Uc
1.0 ○	Median Uc
0.0 ◉	Zero

Value	Parameter	
9	Number of {value, uncertainty} pairs	
0.0	Zero	*Expected Unc*
2.3	Student's t(1-0.95)	
83.6	BS-mean	*Location*
1.6	S(BS-mean)	*Dispersion*
3.7	~U95(population): dispersion*ts	
1.2	~U95(location): dispersion*ts/sqrt(N)	
1.9	%RSD	

{Uwt-mean, Suwt}



Location

Value	Parameter
83.5 ○	Mean
82.5 ◉	Uwt-mean
82.9 ○	MPwt-mean
83.4 ○	Median
83.2 ○	Shorth
83.3 ○	MM-median
82.9 ○	MM-sh/mid
82.9 ○	MM-sh/med
82.4 ○	MM-mode
83.6 ○	BS-mean
83.5 ○	H15
83.5 ○	A15

82.4 Minimum
83.6 Maximum

83.5 KCRV
0.8 KCRU

Dispersion

Value	Parameter
0.8 ○	S & ExU
0.4 ◉	Suwt & ExU
0.3 ○	SMPwt & ExU
0.7 ○	MADe & ExU
0.7 ○	IQR & ExU
0.7 ○	S(Shorth) & ExU
1.3 ○	S(MM-median)
1.0 ○	S(MM-shorth)
1.6 ○	S(BS-mean)
1.5 ○	S(BS-SD)
0.9 ○	S(H15) & ExU

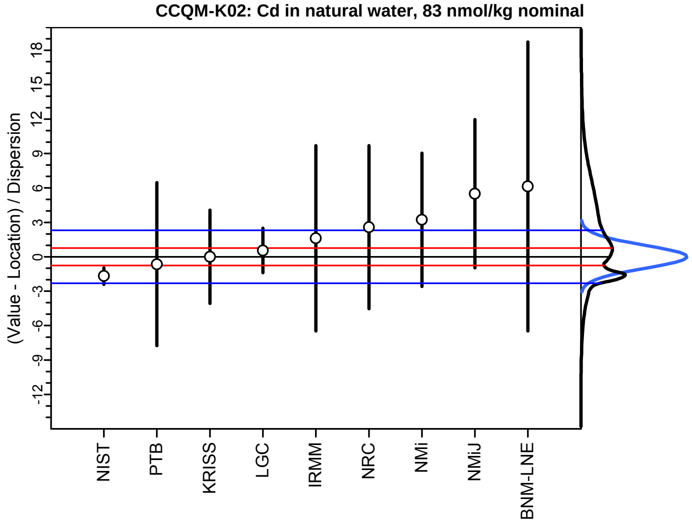
0.3 Minimum
1.6 Maximum

Expected Uncertainty

Value	Parameter
1.1 ○	Pooled Uc
1.0 ○	Median Uc
0.0 ◉	Zero

Value	Parameter	
9	Number of {value, uncertainty} pairs	
0.0	Zero	*Expected Unc*
2.3	Student's t(1-0.95)	
82.5	Uwt-mean	*Location*
0.4	Suwt & ExU	*Dispersion*
0.9	~U95(population): dispersion*ts	
0.3	~U95(location): dispersion*ts/sqrt(N)	
0.5	%RSD	

{MPwt-mean, SMPwt}



Location

Value	Parameter
83.5	○ Mean
82.5	○ Uwt-mean
82.9	◉ MPwt-mean
83.4	○ Median
83.2	○ Shorth
83.3	○ MM-median
82.9	○ MM-sh/mid
82.9	○ MM-sh/med
82.4	○ MM-mode
83.6	○ BS-mean
83.5	○ H15
83.5	○ A15

82.4 Minimum
83.6 Maximum

83.5 KCRV
0.8 KCRU

Dispersion

Value	Parameter
0.8	○ S & ExU
0.4	○ Suwt & ExU
0.3	◉ SMPwt & ExU
0.7	○ MADe & ExU
0.7	○ IQR & ExU
0.7	○ S(Shorth) & ExU
1.3	○ S(MM-median)
1.0	○ S(MM-shorth)
1.6	○ S(BS-mean)
1.5	○ S(BS-SD)
0.9	○ S(H15) & ExU

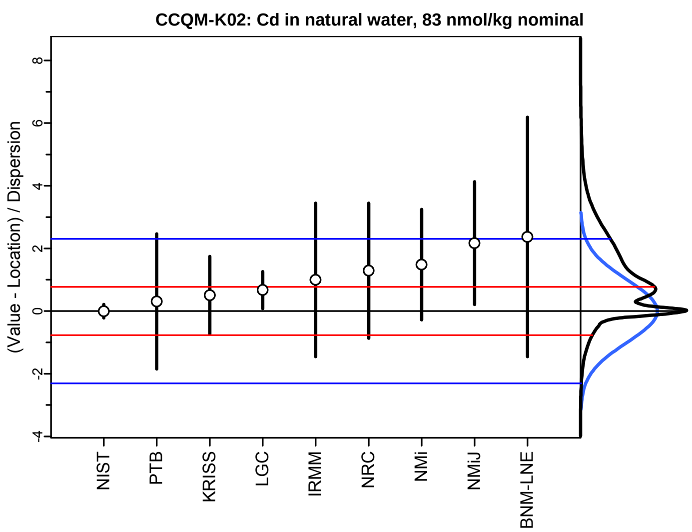
0.3 Minimum
1.6 Maximum

Expected Uncertainty

Value	Parameter
1.1	○ Pooled Uc
1.0	○ Median Uc
0.0	◉ Zero

Value	Parameter	
9	Number of {value, uncertainty} pairs	
0.0	Zero	*Expected Unc*
2.3	Student's t(1-0.95)	
82.9	MPwt-mean	*Location*
0.3	SMPwt & ExU	*Dispersion*
0.7	~U95(population): dispersion*ts	
0.2	~U95(location): dispersion*ts/sqrt(N)	
0.4	%RSD	

{MM-mode, S(MM-shorth)}



Location

Value	Parameter
83.5	○ Mean
82.5	○ Uwt-mean
82.9	○ MPwt-mean
83.4	○ Median
83.2	○ Shorth
83.3	○ MM-median
82.9	○ MM-sh/mid
82.9	○ MM-sh/med
82.4	◉ MM-mode
83.6	○ BS-mean
83.5	○ H15
83.5	○ A15

82.4 Minimum
83.6 Maximum

83.5 KCRV
0.8 KCRU

Dispersion

Value	Parameter
0.8	○ S & ExU
0.4	○ Suwt & ExU
0.3	○ SMPwt & ExU
0.7	○ MADe & ExU
0.7	○ IQR & ExU
0.7	○ S(Shorth) & ExU
1.3	○ S(MM-median)
1.0	◉ S(MM-shorth)
1.6	○ S(BS-mean)
1.5	○ S(BS-SD)
0.9	○ S(H15) & ExU

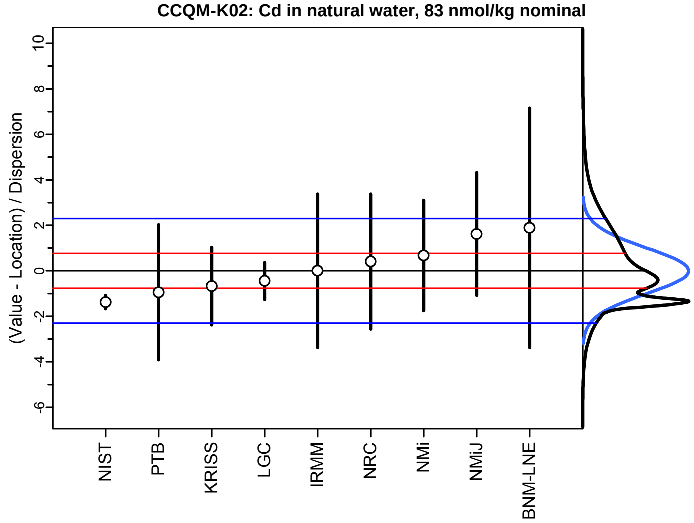
0.3 Minimum
1.6 Maximum

Expected Uncertainty

Value	Parameter
1.1	○ Pooled Uc
1.0	○ Median Uc
0.0	◉ Zero

Value	Parameter	
9	Number of {value, uncertainty} pairs	
0.0	Zero	*Expected Unc*
2.3	Student's t(1-0.95)	
82.4	MM-mode	*Location*
1.0	S(MM-shorth)	*Dispersion*
2.4	~U95(population): dispersion*ts	
0.8	~U95(location): dispersion*ts/sqrt(N)	
1.2	%RSD	

{Median, MADe}



Location

Value		Parameter
83.5	○	Mean
82.5	○	Uwt-mean
82.9	○	MPwt-mean
83.4	◉	Median
83.2	○	Shorth
83.3	○	MM-median
82.9	○	MM-sh/mid
82.9	○	MM-sh/med
82.4	○	MM-mode
83.6	○	BS-mean
83.5	○	H15
83.5	○	A15

82.4 Minimum
83.6 Maximum

83.5 KCRV
0.8 KCRU

Dispersion

Value		Parameter
0.8	○	S & ExU
0.4	○	Suwt & ExU
0.3	○	SMPwt & ExU
0.7	◉	MADe & ExU
0.7	○	IQR & ExU
0.7	○	S(Shorth) & ExU
1.3	○	S(MM-median)
1.0	○	S(MM-shorth)
1.6	○	S(BS-mean)
1.5	○	S(BS-SD)
0.9	○	S(H15) & ExU

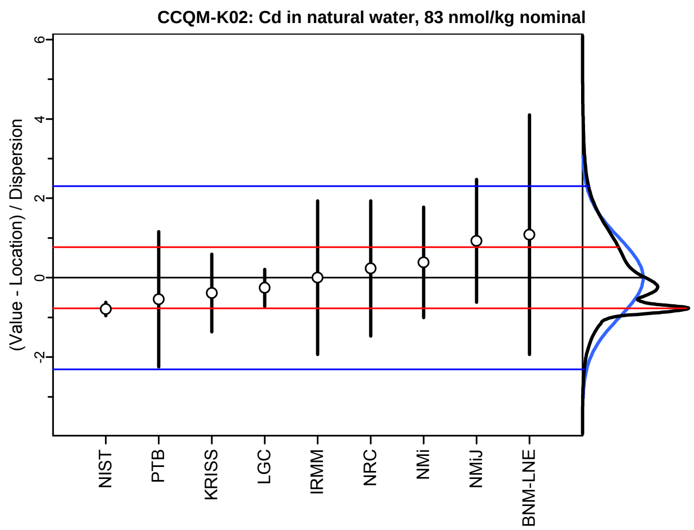
0.3 Minimum
1.6 Maximum

Expected Uncertainty

Value		Parameter
1.1	○	Pooled Uc
1.0	○	Median Uc
0.0	◉	Zero

Value	Parameter	
9	Number of {value, uncertainty} pairs	
0.0	Zero	*Expected Unc*
2.3	Student's t(1-0.95)	
83.4	Median	*Location*
0.7	MADe & ExU	*Dispersion*
1.7	~U95(population): dispersion*ts	
0.6	~U95(location): dispersion*ts/sqrt(N)	
0.9	%RSD	

{Median, MADe+}



Location

Value		Parameter
83.5	○	Mean
82.5	○	Uwt-mean
82.9	○	MPwt-mean
83.4	◉	Median
83.2	○	Shorth
83.3	○	MM-median
82.9	○	MM-sh/mid
82.9	○	MM-sh/med
82.4	○	MM-mode
83.6	○	BS-mean
83.5	○	H15
83.5	○	A15

82.4 Minimum
83.6 Maximum

83.5 KCRV
0.8 KCRU

Dispersion

Value		Parameter
1.3	○	S & ExU
1.1	○	Suwt & ExU
1.1	○	SMPwt & ExU
1.3	◉	MADe & ExU
1.3	○	IQR & ExU
1.3	○	S(Shorth) & ExU
1.3	○	S(MM-median)
1.0	○	S(MM-shorth)
1.6	○	S(BS-mean)
1.5	○	S(BS-SD)
1.4	○	S(H15) & ExU

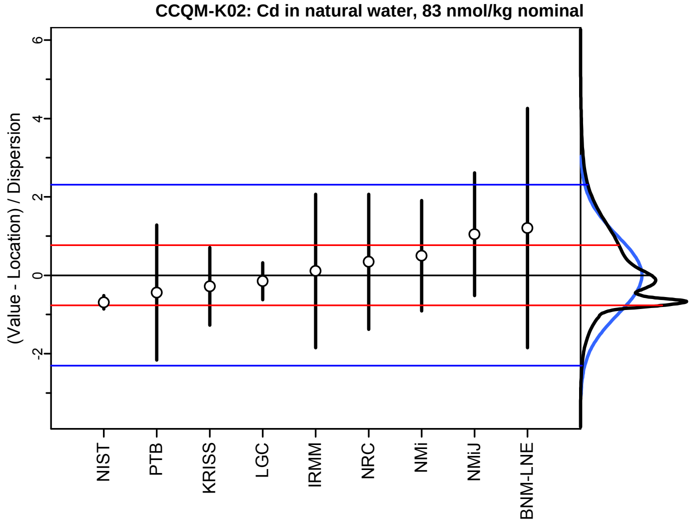
1.0 Minimum
1.6 Maximum

Expected Uncertainty

Value		Parameter
1.1	◉	Pooled Uc
1.0	○	Median Uc
0.0	○	Zero

Value	Parameter	
9	Number of {value, uncertainty} pairs	
1.1	Pooled Uc	*Expected Unc*
2.3	Student's t(1-0.95)	
83.4	Median	*Location*
1.3	MADe & ExU	*Dispersion*
3.0	~U95(population): dispersion*ts	
1.0	~U95(location): dispersion*ts/sqrt(N)	
1.5	%RSD	

{MM-median, S(MM-median)}



Location

Value		Parameter
83.5	○	Mean
82.5	○	Uwt-mean
82.9	○	MPwt-mean
83.4	○	Median
83.2	○	Shorth
83.3	◉	MM-median
82.9	○	MM-sh/mid
82.9	○	MM-sh/med
82.4	○	MM-mode
83.6	○	BS-mean
83.5	○	H15
83.5	○	A15

82.4 Minimum
83.6 Maximum

83.5 KCRV
0.8 KCRU

Dispersion

Value		Parameter
0.8	○	S & ExU
0.4	○	Suwt & ExU
0.3	○	SMPwt & ExU
0.7	○	MADe & ExU
0.7	○	IQR & ExU
0.7	○	S(Shorth) & ExU
1.3	◉	S(MM-median)
1.0	○	S(MM-shorth)
1.6	○	S(BS-mean)
1.5	○	S(BS-SD)
0.9	○	S(H15) & ExU

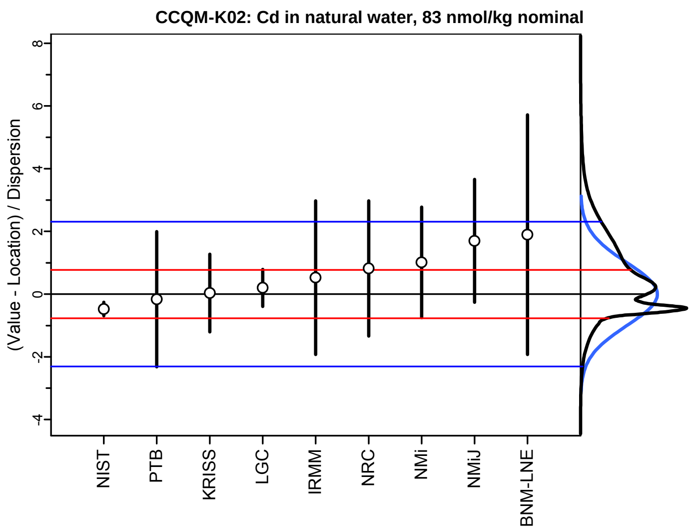
0.3 Minimum
1.6 Maximum

Expected Uncertainty

Value		Parameter
1.1	○	Pooled Uc
1.0	○	Median Uc
0.0	◉	Zero

Value	Parameter	
9	Number of {value, uncertainty} pairs	
0.0	Zero	*Expected Unc*
2.3	Student's t(1-0.95)	
83.3	MM-median	*Location*
1.3	S(MM-median)	*Dispersion*
2.9	~U95(population): dispersion*ts	
1.0	~U95(location): dispersion*ts/sqrt(N)	
1.5	%RSD	

{MM-shorth, S(MM-shorth)}



Location

Value		Parameter
83.5	○	Mean
82.5	○	Uwt-mean
82.9	○	MPwt-mean
83.4	○	Median
83.2	○	Shorth
83.3	○	MM-median
82.9	○	MM-sh/mid
82.9	◉	MM-sh/med
82.4	○	MM-mode
83.6	○	BS-mean
83.5	○	H15
83.5	○	A15

82.4 Minimum
83.6 Maximum

83.5 KCRV
0.8 KCRU

Dispersion

Value		Parameter
0.8	○	S & ExU
0.4	○	Suwt & ExU
0.3	○	SMPwt & ExU
0.7	○	MADe & ExU
0.7	○	IQR & ExU
0.7	○	S(Shorth) & ExU
1.3	○	S(MM-median)
1.0	◉	S(MM-shorth)
1.6	○	S(BS-mean)
1.5	○	S(BS-SD)
0.9	○	S(H15) & ExU

0.3 Minimum
1.6 Maximum

Expected Uncertainty

Value		Parameter
1.1	○	Pooled Uc
1.0	○	Median Uc
0.0	◉	Zero

Value	Parameter	
9	Number of {value, uncertainty} pairs	
0.0	Zero	*Expected Unc*
2.3	Student's t(1-0.95)	
82.9	MM-sh/med	*Location*
1.0	S(MM-shorth)	*Dispersion*
2.4	~U95(population): dispersion*ts	
0.8	~U95(location): dispersion*ts/sqrt(N)	
1.2	%RSD	