Companion Document

Treatment of Elemental Mercury Gas Generator Certification Data and Calculation of Uncertainty

in Support of the

Interim EPA Traceability Protocol for Qualification and Certification of Elemental Mercury Gas Generators

The interim elemental mercury (Hg) traceability gas protocol[1] has been developed with full awareness that most of the calculations, especially the complex uncertainty calculations, would be performed using a computer spreadsheet. In fact, a set of example spreadsheets has been developed and provided as supplemental material to this interim protocol.

This Companion Document provides support and clarification for users of the interim elemental Hg gas traceability protocol and associated spreadsheets. It includes equations and example calculations to facilitate protocol implementation. Section A presents the methodology and equations that are used to calculate the expanded uncertainty of gas generator certifications performed under the protocol. Section B shows detailed example calculations of certification data and uncertainty.

Section A – Certification Uncertainty

For the process of assigning concentration values to candidate generator setpoints, five sources of uncertainty have been identified:

- Calibration Linearity
- Measurement Stability
- Repeatability
- Reproducibility
- Reference Uncertainty

The first three uncertainty components apply specifically to a bracketing procedure. The other two components apply to the overall process of transferring traceability from the reference gas(es) to the candidate.

A.1 Calibration Linearity

[Reserved]

A.2 Measurement Stability

The measurement stability uncertainty calculation applies a statistical test to the assumptions that are built into the bracketing technique (e.g., no non-linear drift, wavering responses, or excessively "noisy" measurements). This uncertainty component is quantified in terms of the residual error estimate for these two lines:

$$SE_{\mathrm{Cand}} = \sqrt{\frac{\left[\sum\left(\mathbf{m}_{\mathrm{Cand}} - \overline{\mathbf{m}_{\mathrm{Cand}}}\right)^2 - \frac{\left[\sum\left(\mathbf{m}_{\mathrm{Cand}} - \overline{\mathbf{m}_{\mathrm{Cand}}}\right)\left(t_{\mathrm{Cand}} - \overline{t_{\mathrm{Cand}}}\right)\right]^2}{\sum\left(t_{\mathrm{Cand}} - \overline{t_{\mathrm{Cand}}}\right)^2}\right]}{(\mathbf{n}_{\mathrm{Cand}} - 2)}}$$

$$SE_{\mathrm{Ref}} = \sqrt{\frac{\left[\sum\left(\mathbf{m}_{\mathrm{Ref}} - \overline{\mathbf{m}_{\mathrm{Ref}}}\right)^2 - \frac{\left[\sum\left(\mathbf{m}_{\mathrm{Ref}} - \overline{\mathbf{m}_{\mathrm{Ref}}}\right)\left(t_{\mathrm{Ref}} - \overline{t_{\mathrm{Ref}}}\right)\right]^2}{\sum\left(t_{\mathrm{Ref}} - \overline{t_{\mathrm{Ref}}}\right)^2}\right]}{(\mathbf{n}_{\mathrm{Ref}} - 2)}}$$

where t_{Cand} and t_{Ref} are the timestamps associated with the candidate and reference measurements. The t units and the $t = 0$ origin can be anything, as long as they are consistent within each standard error calculation. The bold $\mathbf{m}$ and $\mathbf{n}$ values refer to all of the points of the regression lines. The standard errors for these lines are used to calculate the standard error of the mean of each individual measurement interval.

$$SE(\overline{\mathrm{m}}_{\mathrm{Ref\text{-}B}}) = \frac{SE_{\mathrm{Ref}}}{\sqrt{\mathrm{n}_{\mathrm{Ref\text{-}B}}}} \qquad SE(\overline{\mathrm{m}}_{\mathrm{Cand}}) = \frac{SE_{\mathrm{Cand}}}{\sqrt{\mathrm{n}_{\mathrm{Cand}}}} \qquad SE(\overline{\mathrm{m}}_{\mathrm{Ref\text{-}A}}) = \frac{SE_{\mathrm{Ref}}}{\sqrt{\mathrm{n}_{\mathrm{Ref\text{-}A}}}}$$

where $\overline{\mathrm{m}}$ is the mean for just that measurement interval, and n is the number of individual measurements in that interval. The subscript notations (i.e., Ref-B, Cand, and Ref-A) relate to the before-candidate reference measurement, the candidate measurement, and the after-candidate reference measurement (respectively). For each ratio calculation, these standard errors are used to calculate a combined stability uncertainty for this ratio

$$R_{jk} = \frac{\overline{\mathrm{m}}_{\mathrm{Cand}}}{\frac{t_{\mathrm{Ref\text{-}A}} - t_{\mathrm{Cand}}}{t_{\mathrm{Ref-A}} - t_{\mathrm{Ref\text{-}B}}} \cdot \overline{\mathrm{m}}_{\mathrm{Ref\text{-}B}} + \frac{t_{\mathrm{Cand}} - t_{\mathrm{Ref\text{-}B}}}{t_{\mathrm{Ref-A}} - t_{\mathrm{Ref\text{-}B}}} \cdot \overline{\mathrm{m}}_{\mathrm{Ref\text{-}A}}}$$

$$u_{1\sigma}(R_{jk})_{stability} = R_{jk} \times \sqrt{\left(\frac{SE(\overline{\mathrm{m}}_{\mathrm{Ref\text{-}B}})}{\frac{t_{\mathrm{Ref\text{-}A}} - t_{\mathrm{Reb\text{-}B}}}{t_{\mathrm{Ref-A}} - t_{\mathrm{Cand}}} \times \overline{\mathrm{m}}_{\mathrm{Ref\text{-}B}}}\right)^2 + \left(\frac{SE(\overline{\mathrm{m}}_{\mathrm{Cand}})}{\overline{\mathrm{m}}_{\mathrm{Cand}}}\right)^2 + \left(\frac{SE(\overline{\mathrm{m}}_{\mathrm{Ref\text{-}A}})}{\frac{t_{\mathrm{Ref\text{-}A}} - t_{\mathrm{Reb\text{-}B}}}{t_{\mathrm{Cand}} - t_{\mathrm{Ref\text{-}B}}} \times \overline{\mathrm{m}}_{\mathrm{Ref\text{-}A}}}\right)^2}$$

where R_{jk} is the measured ratio for the k^{th} individual bracket within the j^{th} contiguous set of brackets.

Measurement stability is one of three uncertainty components (the first being detector linearity) that relate only to a single bracketing procedure, with average ratio

$$\overline{R}_j = \frac{\sum_{k=1}^{K} R_{jk}}{K}$$

and measurement stability uncertainty

$$u_{1\sigma}(\overline{R}_j)_{stability} = \sqrt{\frac{1}{K^2}\sum_{k=1}^{K} u^2(R_{jk})_{stability}}$$

where K is the number of brackets in this j^{th} set.

Note: Some measurement systems, by design, have more than one measurement "channel" capable of making concurrent independent concentration measurements. If more than one channel is used, all calculations of ratios and uncertainties must be performed independently, up to the point of calculating $\overline{R}_j$ and the bracketing uncertainty $u(\overline{R}_j)$ for the set.

A.3 Repeatability

The most straightforward way to calculate repeatability is a simple standard deviation of the individual bracket ratios.

$$s\left(R_{jk}\right) = \sqrt{\frac{\sum_{k=1}^{K}\left(R_{jk} - \overline{R}_j\right)^2}{(K-1)}}$$

This approach, however, creates the possibility of "double-counting" where measurement instability contributes to imprecision. To allow for this, the experiments are structured as a 3-level nested design (see chapter 2 of NIST's "Engineering Statistical Handbook" http://www.sbtionline.com/nist/mpc/mpc.htm), as follows:

Level 1: Individual ratio uncertainty
Level 2: Repeatability within set (sometimes called "within day" uncertainty)
Level 3: Set reproducibility (sometimes called "between day" uncertainty)

In a nested experimental design, standard errors are pooled at each level, and the level-specific contributions to uncertainty are computed from these pooled standard errors. Starting with the pooled level-1 standard error

$$s_1 = \sqrt{\frac{1}{K}\sum_{k=1}^{K} u^2\left(R_{jk}\right)_{stability}}$$

and the level-2 standard deviation

$$s_2 = \sqrt{\frac{\sum_{k=1}^{K}\left(R_{jk} - \overline{R}_j\right)^2}{(K-1)}}$$

The repeatability uncertainty of the ratio measurements is computed using

$$u_{1\sigma}\left(R_j\right)_{\text{repeatability}} = \sqrt{Max\left[0\,,\; s_2^2 - \frac{1}{K}\times s_1^2\right]} = \sqrt{Max\left[0\,,\; \frac{\sum_{k=1}^{K}\left(\mathrm{R_{jk}} - \overline{\mathrm{R}}_\mathrm{j}\right)^2}{(\mathrm{K}-1)} - \frac{\sum_{k=1}^{K}\mathrm{u}^2\left(\mathrm{R_{jk}}\right)_{stability}}{\mathrm{L}\times\mathrm{K}}\right]}$$

where L represents the average number of individual detector readings used in each measurement average and ratio determination

$$L = \frac{1}{K}\sum_{k=1}^{K}\left(\mathrm{n_{Ref\text{-}B,k}} + \mathrm{n_{Cand,k}} + \mathrm{n_{Ref\text{-}A,k}}\right)$$

The uncertainty of the set-average ratio is

$$u_{1\sigma}\left(\overline{R}_j\right)_{\text{repeatability}} = \frac{1}{\sqrt{K}}u_{1\sigma}\left(R_j\right)_{\text{repeatability}} = \frac{1}{\sqrt{K}}\times\sqrt{Max\left[0\,,\; \frac{\sum_{k=1}^{K}\left(\mathrm{R_{jk}} - \overline{\mathrm{R}}_\mathrm{j}\right)^2}{(\mathrm{K}-1)} - \frac{\sum_{k=1}^{K}\mathrm{u}^2\left(\mathrm{R_{jk}}\right)_{stability}}{\mathrm{L}\times\mathrm{K}}\right]}$$

The "Max" function holds this term at zero when short-term precision uncertainty is dominated by measurement instability.

A.4 Reproducibility

Reproducibility is an uncertainty term applied to the calculated candidate concentrations ($C_{\mathrm{Cand,j}} = \overline{R}_j \times C_{\mathrm{Ref}}$). It may be estimated differently depending on what type of generator is being certified (i.e., Vendor-Prime, Field-Reference , or User generator) and how many times the bracketing procedure is repeated.

A.4.1 The Statistical Approach

A 3-level nested design[2] assumes that the ratio measurement is done K>1 times on each of J>1 days. Using the same technique as for the repeatability uncertainty

$$u_{1\sigma}\left(C_{\mathrm{Cand}}\right)_{\mathrm{reproducibility}} = \sqrt{Max\left[0, s_3^2 - \frac{1}{J} \times s_{2,C_{\mathrm{Cand}}}^2\right]}$$

where

$$s_3 = \sqrt{\frac{1}{J-1} \times \sum_{j=1}^{J} \left(C_{\mathrm{Cand,j}} - \overline{C_{\mathrm{Cand}}}\right)^2}$$

and s_2 values (from A.3) are converted to C_{Cand} errors and pooled as follows

$$s_{2,C_{cand}} = \sqrt{\frac{1}{\sum K_j - J} \times \sum_{j=1}^{J} \left[(K_j - 1) \times \left(C_{\mathrm{Ref},j} \times s_{2,j}\right)^2\right]}$$

As with the repeatability uncertainty, the "Max" function holds this term at zero when the deviation among sets is dominated by short-term effects. The reproducibility uncertainty of the mean candidate concentration is

$$u_{1\sigma}\left(\overline{C_{\mathrm{Cand}}}\right)_{\mathrm{reproducibility}} = \frac{1}{\sqrt{J}} \sqrt{Max\left[0, s_3^2 - \frac{1}{K} \times s_{2,C_{\mathrm{Cand}}}^2\right]} \quad \text{for} \quad \overline{C_{\mathrm{Cand}}} = \sum\left(\overline{R}_j \times C_{\mathrm{Ref,j}}\right) / J$$

A.4.2 The Bound on Bias (BOB) Approach

When the number of bracketing data sets are few (≤5 sets), a shorthand approach, described in detail elsewhere[3], can be used to compute the reproducibility standard error. This is necessary because small sample sizes don't provide a lot of information about population distribution. The BOB approach assumes a uniform distribution (as opposed to a normal distribution), and sets the bounds at the highest and lowest C_{Cand} values among the data. The estimated standard deviation for a uniform distribution is

$$s_{uniform} = \frac{a}{\sqrt{3}}$$

where a is half the width of the distribution. Applying this formula to all the bracketing data yields

$$u_{1\sigma}\left(\overline{C_{\text{Cand}}}\right)_{\text{reproducibility}} = \frac{Max(C_{\text{Cand}}) - Min(C_{\text{Cand}})}{2\times\sqrt{3}} = \frac{Max(C_{\text{Cand}}) - Min(C_{\text{Cand}})}{\sqrt{12}}$$

A.4.3 The Type-B Uncertainty Approach

For User generators, where there are no "downstream" certifications to inherit uncertainty, a "default" value may be used instead of repeating the bracketing procedure (a particularly useful alternative when the certification is performed in the field). This value, based on what is routinely achievable by commercially available systems, is

$$u_{1\sigma}\left(\overline{C_{\text{Cand}}}\right)_{\text{reproducibility}} = 0.005\times\overline{C_{\text{Cand}}}$$

A.5 Combined Uncertainty

Once calculated, all of the uncertainties are combined using propagation of error. First, the combined uncertainty of each bracketing procedure is calculated

$$u_{1\sigma}\left(\frac{C_{\text{Cand}}}{C_{\text{Ref}}}\right)_j = u_{1\sigma}\left(\overline{R}_j\right) = \sqrt{\left[u\left(\overline{R}_j\right)_{\text{linearity}} \underset{default}{\overset{assumed}{\approx 0}}\right]^2 + u_{1\sigma}^2\left(\overline{R}_j\right)_{\text{stability}} + u_{1\sigma}^2\left(\overline{R}_j\right)_{\text{repeatability}}}$$

Then, these uncertainties are combined into a single comparison uncertainty

$$u_{1\sigma}\left(\overline{C_{\text{Cand}}}\right)_{comparison} = \sqrt{\frac{1}{J^2}\sum_{j=1}^{J}\left[C_{\text{Ref},j}\times u_{1\sigma}\left(\overline{R}_j\right)\right]^2}$$

The reference gas uncertainty component is calculated based on its reported value and the average comparison ratio

$$u_{1\sigma}\left(\overline{C_{\text{Cand}}}\right)_{\text{Reference}} = \frac{\overline{C_{\text{Cand}}}}{C_{\text{Ref}}}\times u_{1\sigma}\left(C_{\text{Ref}}\right)_{\text{certification}}$$

The overall combined uncertainty for the candidate gas certification is

$$u_{1\sigma}\left(\overline{C_{\text{Cand}}}\right)_{combined} = \sqrt{u_{1\sigma}^2\left(\overline{C_{\text{Cand}}}\right)_{\text{Comparison}} + u_{1\sigma}^2\left(\overline{C_{\text{Cand}}}\right)_{\text{Reference}} + u_{1\sigma}^2\left(\overline{C_{\text{Cand}}}\right)_{\text{reproducib ility}}}$$

and the expanded certification uncertainty (k=2) is

$$u_{1\sigma}\left(\overline{C_{\text{Cand}}}\right)_{\exp anded} = k\times u_{1\sigma}\left(\overline{C_{\text{Cand}}}\right)_{combined} = 2\times u_{1\sigma}\left(\overline{C_{\text{Cand}}}\right)_{combined}$$

Section B – Example Calculations

These example calculations show how bracketing data are used to certify candidate generators, determining concentration and uncertainty. Bracketing data are collected as a series of Hg^0 detector readings over time. The trace shown in Figure 1 represents detector readings from one candidate generator and a reference generator at one setpoint (nominal 10 µg/m^3), along with pre- and post-test zero readings.

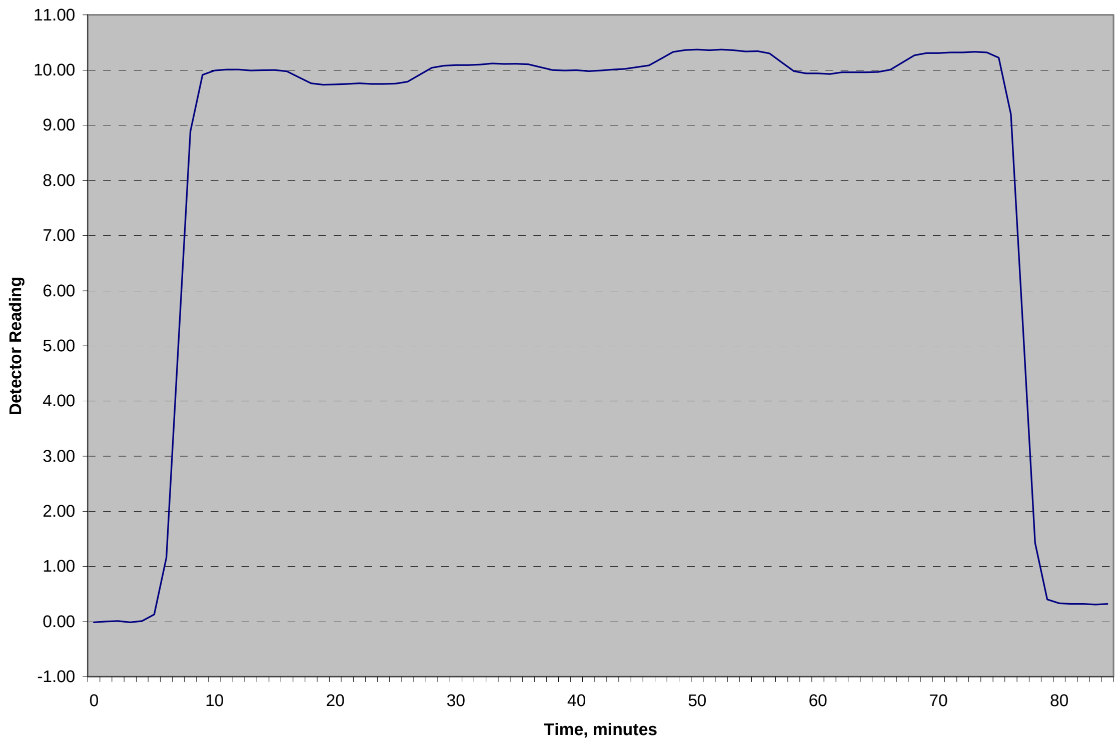

Figure 1. Example Bracketing Data Trace

Table B-1 parses out 5 readings for each measurement interval, and establishes a timestamp to be used in the calculations. In this example, there are two zero measurement intervals (before and after the bracketing test), three candidate generator measurement intervals, and four reference generator measurement intervals. For consistency with the example formulae shown in Section 5.2 of EPA's interim elemental Hg traceability protocol[1], these data will represent the High setpoint certification of a user generator (U_H) using a field-reference generator (FR_H^*).

Table B-1. Parsed Bracketing Data Example.

Detector Measuring	Time (min)	Reading #1	Reading #2	Reading #3	Reading #4	Reading #5	Measurement Average
Zero	0	-0.01	0.00	0.00	0.00	0.01	0.00
Reference	10	9.99	9.97	10.00	10.02	10.02	10.00
Candidate	20	9.84	9.86	9.84	9.86	9.85	9.85
Reference	30	10.20	10.19	10.21	10.20	10.21	10.20
Candidate	40	9.95	9.95	9.97	9.96	9.97	9.96
Reference	50	10.26	10.25	10.23	10.25	10.26	10.25
Candidate	60	9.98	9.99	10.00	10.01	10.01	10.00
Reference	70	10.41	10.39	10.41	10.40	10.40	10.40
Zero	80	0.32	0.33	0.32	0.30	0.33	0.32

B.1 Zero Correction

The zero offset is calculated using time-interpolation of the pre- and post-test zero measurements, based on Equation 3 in Section 5.2 of the interim elemental Hg traceability protocol[1]. Since the individual readings are equally spaced, and the measurement averages are also equally spaced and represent the same number of readings (5 in this case), the zero offset formula can be applied to either these readings or the measurement averages (the two approaches are mathematically identical). For this example, the zero correction will be applied to the measurement averages from Table B-1, so for the first reference measurement (Time $t_{\text{Ref-B}} = 10$), the zero offset is calculated as

$$Z_{\text{Ref-B1}} = Z_1 + \left[\left(t_{\text{Ref-B1}} - t_1\right) \frac{\left(Z_1 - Z_2\right)}{\left(t_1 - t_2\right)} \right] = 0 + \left[\left(10 - 0\right) \frac{\left(0 - 0.32\right)}{\left(0 - 80\right)} \right] = 0.04$$

and the zero-corrected measurement average is

$$V_c *_{\text{Ref-B1}} = V *_{\text{Ref-B1}} - Z_{\text{Ref-B1}} = 10.00 - 0.04 = 9.96$$

Table B-2 shows all of the zero correction values for the example dataset. These data are used for calculating bracketing ratios when the measurement system does not perform automatic background correction for every reading.

Table B-2. Zero Correction of Example Data.

Time (min)	Zero Response	Interpolated Zero	FR_i* Response	FR_c* Zero-Cor.	U_i Response	U_c Zero-Cor.
0	0.00					
10		0.04	10.00	9.96		
20		0.08			9.85	9.77
30		0.12	10.20	10.08		
40		0.16			9.96	9.80
50		0.20	10.25	10.05		
60		0.24			10.00	9.76
70		0.28	10.40	10.12		
80	0.32					

B.2 Output Ratio

The output ratio is calculated from the average candidate measurement and the time-interpolated average reference measurement. Using the formula Equation A-1 and the first bracket of zero-corrected example data from Table B-2

$$Ratio = \frac{\overline{m}_{Cand}}{\frac{t_{Ref\text{-}A} - t_{Cand}}{t_{Ref-A} - t_{Ref\text{-}B}} \cdot \overline{m}_{Ref\text{-}B} + \frac{t_{Cand} - t_{Ref\text{-}B}}{t_{Ref-A} - t_{Ref\text{-}B}} \cdot \overline{m}_{Ref\text{-}A}} = \frac{9.77}{\frac{30-20}{30-10} \cdot 9.96 + \frac{20-10}{30-10} \cdot 10.08} = 0.975$$

Since the measurements for this example are equally spaced in time, the formula gives the same results as Equation 2 in Section 5. Table B-3 shows the ratios for this example.

Table B-3. Example Output Ratio Data.

Time (min)	FR_i* Response	FR_c* Zero-Cor.	U_i Response	U_c Zero-Cor.	Output Ratio
10	10.00	9.96			
20			9.85	9.77	0.975
30	10.20	10.08			
40			9.96	9.80	0.974
50	10.25	10.05			
60			10.00	9.76	0.968
70	10.40	10.12			
				Average Ratio	0.972

B.3 Ratio RSD

The first check of the validity of a bracketing test is the RSD among the calculated ratios. Using example data from Table B-3 and Equation 5 from Section 5.2 of the interim elemental Hg traceability protocol[1].

$$RSD = \frac{\sqrt{\frac{\sum_{i=1}^{n}\left(x_i - \bar{x}\right)^2}{n-1}}}{\bar{x}} \ x \ 100\% = \frac{\sqrt{\frac{(.975-.972)^2+(.974-.972)^2+(.968-.972)^2}{3-1}}}{.972} \ x \ 100\% = 0.4\%$$

B.4 Certified Concentration of User Generator (Single Bracketing Option)

If the certified user generator concentration is to be based on a single bracketing procedure, the concentration is calculated from the average bracketing ratio and the certified reference concentration. For the example of Table B-3, where the certified field reference generator output (XFR_H^*) is assumed to be 10.0 µg/m^3, the certified user generator concentration is calculated as

$$Y_H = XFR_H^* \times Ratio_{Average} = 10.0\,\mu g/m^3 \times 0.972 = 9.72\,\mu g/m^3$$

B.5 Detector Linearity Uncertainty

[Reserved]

B.6 Measurement Stability Uncertainty

The measurement stability uncertainty calculation uses the individual readings recorded during the bracketing procedure, prior to any averaging or zero correction. Using the example data from Table B-1, the readings and timestamps are separated into two datasets, one for all of the reference generator readings and one for all of the candidate generator readings. For each of these datasets, statistical “standard error of estimate” calculation is performed with the timestamps as the independent variable. Table B-4 shows some of the intermediate calculations for the example candidate readings.

Table B-4. Intermediate Calculations for Measurement Stability

t_{Cand} Time	m_{Cand} Reading	$\left(t_{Cand} - \overline{t_{Cand}}\right)^2$	$\left(m_{Cand} - \overline{m_{Cand}}\right)^2$	$\left(m_{Cand} - \overline{m_{Cand}}\right)\left(t_{Cand} - \overline{t_{Cand}}\right)$
20	9.84	144	0.0087	1.121
21	9.86	121	0.0066	0.895
22	9.84	100	0.0088	0.938
23	9.86	81	0.0060	0.695
24	9.85	64	0.0077	0.700
40	9.95	4	0.0003	-0.032
41	9.95	1	0.0002	-0.012
42	9.97	0	0.0010	0.000
43	9.96	1	0.0007	0.027
44	9.97	4	0.0009	0.060
60	9.98	64	0.0022	0.379
61	9.99	81	0.0026	0.460
62	10.00	100	0.0043	0.656
63	10.01	121	0.0055	0.819
64	10.01	144	0.0061	0.939
Averages		Summations		
42	9.94	4030	0.0616	15.143

These data are entered into the standard error of estimate formula

$$SE_{Cand} = \sqrt{\frac{\left[\sum\left(m_{Cand} - \overline{m_{Cand}}\right)^2 - \frac{\left[\sum\left(m_{Cand} - \overline{m_{Cand}}\right)\left(t_{Cand} - \overline{t_{Cand}}\right)\right]^2}{\sum\left(t_{Cand} - \overline{t_{Cand}}\right)^2}\right]}{(n_{Cand} - 2)}}$$

$$SE_{Cand} = \sqrt{\frac{0.0616 - 15.143^2 / 4030}{(15 - 2)}} = .0190$$

Using all of the reference readings in a similar formula

$$SE_{\text{Ref}} = \sqrt{\frac{\left[\sum\left(m_{\text{Ref}} - \overline{m_{\text{Ref}}}\right)^2 - \frac{\left[\sum\left(m_{\text{Ref}} - \overline{m_{\text{Ref}}}\right)\left(t_{\text{Ref}} - \overline{t_{\text{Ref}}}\right)\right]^2}{\sum\left(t_{\text{Ref}} - \overline{t_{\text{Ref}}}\right)^2}\right]}{(n_{\text{Ref}} - 2)}} = 0.0347$$

These two quantities are used to calculate the standard error for each measurement average that is used to calculate output ratios. For the first bracket in the example, the standard errors are

$$SE\left(\overline{m}_{\text{Ref-B}}\right) = \frac{SE_{\text{Ref}}}{\sqrt{n_{\text{Ref-B}}}} = \frac{0.0347}{\sqrt{5}} = 0.0155$$

$$SE\left(\overline{m}_{\text{Cand}}\right) = \frac{SE_{\text{Cand}}}{\sqrt{n_{\text{Cand}}}} = \frac{0.0190}{\sqrt{5}} = 0.0085$$

$$SE\left(\overline{m}_{\text{Ref-A}}\right) = \frac{SE_{\text{Ref}}}{\sqrt{n_{\text{Ref-A}}}} = \frac{0.0347}{\sqrt{5}} = 0.0155$$

Propagating these errors through the output ratio calculation

$$u_{1\sigma}\left(R_{jk}\right)_{stability} = R_{jk} \times \sqrt{\left(\frac{SE\left(\overline{m}_{\text{Ref-B}}\right)}{\frac{t_{\text{Ref-A}} - t_{\text{Reb-B}}}{t_{\text{Ref-A}} - t_{\text{Cand}}} \times \overline{m}_{\text{Ref-B}}}\right)^2 + \left(\frac{SE\left(\overline{m}_{\text{Cand}}\right)}{\overline{m}_{\text{Cand}}}\right)^2 + \left(\frac{SE\left(\overline{m}_{\text{Ref-A}}\right)}{\frac{t_{\text{Ref-A}} - t_{\text{Reb-B}}}{t_{\text{Cand}} - t_{\text{Ref-B}}} \times \overline{m}_{\text{Ref-A}}}\right)^2}$$

$$u_{1\sigma}\left(R_{1,1}\right)_{stability} = 0.975 \times \sqrt{\left(\frac{0.0155}{2 \times 9.96}\right)^2 + \left(\frac{0.0085}{9.77}\right)^2 + \left(\frac{0.0155}{2 \times 10.08}\right)^2} = .00136$$

Similarly, for the other two ratios in the bracketing procedure

$$u_{1\sigma}\left(R_{1,2}\right)_{stability} = 0.974 \times \sqrt{\left(\frac{0.0155}{2 \times 10.08}\right)^2 + \left(\frac{0.0085}{9.80}\right)^2 + \left(\frac{0.0155}{2 \times 10.05}\right)^2} = .00136$$

$$u_{1\sigma}\left(R_{1,1}\right)_{stability} = 0.968 \times \sqrt{\left(\frac{0.0155}{2 \times 10.05}\right)^2 + \left(\frac{0.0085}{9.76}\right)^2 + \left(\frac{0.0155}{2 \times 10.12}\right)^2} = .00135$$

The measurement stability uncertainty of the ratio average is

$$u_{1\sigma}\left(\overline{R}_1\right)_{stability} = \sqrt{\frac{1}{K^2}\sum_{k=1}^{K} u^2\left(R_{1,k}\right)_{stability}} = \sqrt{\frac{1}{9} \cdot \left[(.00136)^2 + (.00136)^2 + (.00135)^2\right]} = .00078$$

B.7 Repeatability

Using the nested design explained in Section A, above, the repeatability formula is

$$u_{1\sigma}\left(\overline{R}_j\right)_{\text{repeatability}} = \frac{1}{\sqrt{K}}\sqrt{Max\left[0\ ,\ s_2^2 - \frac{1}{L}\times s_1^2\right]} = \frac{1}{\sqrt{K}}\times\sqrt{Max\left[0\ ,\ \frac{\sum_{k=1}^{K}\left(\mathrm{R}_{jk}-\overline{\mathrm{R}}_{j}\right)^2}{(\mathrm{K}-1)} - \frac{\sum_{k=1}^{K}\mathrm{u}^2\left(\mathrm{R}_{jk}\right)_{stability}}{\mathrm{L}\times\mathrm{K}}\right]}$$

For the example data shown in Table B-2, where

$$L = \frac{1}{K}\sum_{k=1}^{K}\left(\mathrm{n}_{\text{Ref-B,k}} + \mathrm{n}_{\text{Cand,k}} + \mathrm{n}_{\text{Ref-A,k}}\right) = \frac{1}{3}\sum_{k=1}^{3}\left(5+5+5\right) = 15$$

$$s_1 = \sqrt{\frac{\left(0.00136\right)^2 + \left(0.00136\right)^2 + \left(0.00135\right)^2}{3}} = 0.00136$$

$$s_2 = \sqrt{\frac{\left(0.975-0.972\right)^2 + \left(0.974-0.972\right)^2 + \left(0.968-0.972\right)^2}{(3-1)}} = 0.00381$$

$$u_{1\sigma}\left(\overline{R}_1\right)_{\text{repeatability}} = \frac{1}{\sqrt{3}}\sqrt{Max\left[0\ ,\ 0.00381^2 - \frac{1}{15}\times 0.00136^2\right]} = 0.00219$$

B.8 Combined Bracketing Uncertainty

For each bracketing comparison, the combined uncertainty is a function of detector linearity, measurement stability, and output ratio repeatability.

$$u_{1\sigma}\left(\overline{R}_j\right) = \sqrt{\left[u\left(\overline{R}_j\right)_{\text{linearity}} \underset{default}{\overset{assumed}{\approx 0}}\right]^2 + u_{1\sigma}^2\left(\overline{R}_j\right)_{\text{stability}} + u_{1\sigma}^2\left(\overline{R}_j\right)_{\text{repeatability}}}$$

From the example uncertainty components calculated above

$$u_{1\sigma}\left(\overline{R}_1\right) = \sqrt{0 + (0.00078)^2 + (0.00219)^2} = 0.00232$$

B.9 Uncertainty of User Generator Certification (Single Bracketing Option)

The single bracketing option uses a default value for bracketing reproducibility, which for the example of Table B-3 is

$$u_{1\sigma}\left(Y_{U_H^*}\right)_{\text{reproducibility}} = 0.005\times Y_{U_H^*} = 0.005\times 9.72\ \mu\text{g/m}^3 = 0.0486\ \mu\text{g/m}^3$$

The combined certification uncertainty is

$$u_{1\sigma}\left(Y_{U_H}^*\right)_{combined} = \sqrt{u_{1\sigma}^2\left(Y_{U_H}^*\right)_{\text{Comparison}} + u_{1\sigma}^2\left(Y_{U_H}^*\right)_{\text{Reference}} + u_{1\sigma}^2\left(Y_{U_H}^*\right)_{\text{reproducibility}}}$$

where the uncertainty of the field reference certification is assumed to be 0.06 µg/m^3

$$u_{1\sigma}\left(Y_{U_H}^*\right)_{\text{Reference}} = \overline{R} \times u_{1\sigma}\left(X_{FR_H}^*\right)_{\text{Certification}} = 0.972 \times 0.06 = 0.058$$

and the comparison uncertainty for the single bracketing option (J=1) is

$$u_{1\sigma}\left(Y_{U_H}^*\right)_{comparison} = \sqrt{\frac{1}{J^2}\sum_{j=1}^{J}\left[X_{FR_H}^* \times u_{1\sigma}\left(\overline{R}_j\right)\right]^2} = 10.0\,\mu\text{g/m}^3 \times 0.00232 = 0.0232\,\mu\text{g/m}^3$$

so the combined certification uncertainty is

$$u_{1\sigma}\left(Y_{U_H}^*\right)_{combined} = \sqrt{(0.0232)^2 + (0.058)^2 + (0.0486)^2} = 0.08\,\mu\text{g/m}^3$$

and the expanded certification uncertainty (k=2) is

$$u_{1\sigma}\left(Y_{U_H}^*\right)_{expanded} = k \times u_{1\sigma}\left(Y_{U_H}^*\right)_{combined} = 2 \times 0.08\,\mu\text{g/m}^3 = 0.16\,\mu\text{g/m}^3$$

Calculating the % uncertainty

$$\frac{u_{1\sigma}\left(Y_{U_H}^*\right)_{expanded}}{Y_{U_H}^*} \times 100\% = \frac{0.16\,\mu\text{g/m}^3}{9.72\,\mu\text{g/m}^3} \times 100\% = 1.6\%$$

As shown, the expanded, combined uncertainty is well within the 5% acceptance criterion of Section 6.4 of the interim elemental Hg gas traceability protocol[1].

B.10 Dual-Bracketing Certification Example

Table B-5 expands the earlier example to include a second bracketing test, as would be done for certifying a vendor-prime or field-reference generator. For this example, zero corrections are performed just as before, time-interpolating between each pair of zero responses. Response ratios are calculated for each bracket (K=3), each set of brackets (J=2), and the overall average ratio.

Table B-4. Dual-Bracketing Example Data.

Time (min)	Zero Response	Interpolated Zero	V_i* Response	V_c* Zero-Cor.	FR_i Response	FR_c Zero-Cor.
0	0.00					
10		0.02	10.00	9.98		
20		0.04			9.85	9.81
30		0.06	10.20	10.14		
40		0.08			9.96	9.88
50		0.10	10.25	10.15		
60		0.12			10.00	9.88
70		0.14	10.40	10.26		
240	0.48					
250		0.49	10.45	9.96		
260		0.50			10.15	9.65
270		0.51	10.50	9.99		
280		0.52			10.20	9.68
290		0.53	10.48	9.95		
300		0.54			10.22	9.68
310		0.55	10.50	9.95		
360	0.60					

$$R_{1,1} = \frac{9.81}{\frac{30-20}{30-10}\cdot 9.98 + \frac{20-10}{30-10}\cdot 10.14} = 0.975 \qquad R_{2,1} = \frac{9.65}{\frac{270-260}{270-250}\cdot 9.96 + \frac{260-250}{270-250}\cdot 9.99} = 0.967$$

$$R_{1,2} = \frac{9.88}{\frac{50-40}{50-30}\cdot 10.14 + \frac{40-30}{50-30}\cdot 10.15} = 0.974 \qquad R_{2,2} = \frac{9.68}{\frac{290-280}{290-270}\cdot 9.99 + \frac{280-270}{290-270}\cdot 9.95} = 0.971$$

$$R_{1,3} = \frac{9.88}{\frac{70-60}{70-50}\cdot 10.15 + \frac{60-50}{70-50}\cdot 10.26} = 0.968 \qquad R_{2,3} = \frac{9.68}{\frac{310-300}{310-290}\cdot 9.95 + \frac{300-290}{310-290}\cdot 9.95} = 0.973$$

$$\overline{R}_1 = \frac{0.975 + 0.974 + 0.968}{3} = 0.972 \qquad \overline{R}_2 = \frac{0.967 + 0.971 + 0.973}{3} = 0.970$$

$$RSD_1 = \frac{\sqrt{\dfrac{(.975-.972)^2 + (.974-.972)^2 + (.968-.972)^2}{3-1}}}{.972} \; x \; 100\% = 0.4\%$$

$$RSD_2 = \frac{\sqrt{\dfrac{(.967-.970)^2 + (.971-.970)^2 + (.973-.970)^2}{3-1}}}{.970} \; x \; 100\% = 0.3\%$$

$$\overline{R} = \frac{0.972 + 0.970}{2} = 0.971$$

Assuming a certified vendor-prime concentration of $Xv_H^* = 10.0\ \mu g/m^3$, the certified field-reference concentration is

$$Y_{FR}{}_H^* = 10.0\ \mu g/m^3 \times 0.971 = 9.71\ \mu g/m^3$$

B.11 Dual-Bracketing Uncertainty (BOB Reproducibility)

For each contiguous set of brackets, the combined bracketing uncertainty is calculated as in sections B-5 through B-8. Table B-6 shows uncertainty data for a dual-bracketing example.

Table B-5. Dual-Bracketing Example Uncertainty Components

	Bracketing Test Number (j)	
Uncertainty Component	**1**	**2**
Detector Linearity, $u_{1\sigma}\left(\overline{R}_j\right)_{linearity}$	0	0
Measurement Stability, $u_{1\sigma}\left(\overline{R}_j\right)_{stability}$	0.00078	0.00082
Repeatability, $u_{1\sigma}\left(\overline{R}_j\right)_{repeatability}$	0.00219	0.00175
Combined Bracketing, $u_{1\sigma}\left(\overline{R}_j\right)$	0.00232	0.00193

The comparison uncertainty, based on all bracketing tests, is calculated by

$$u_{1\sigma}\left(Y_{FR}{}_H^*\right)_{comparison} = \sqrt{\frac{1}{J^2}\sum_{j=1}^{J}\left[C_{\mathrm{Ref},j} \times u_{1\sigma}\left(\overline{R}_j\right)\right]^2} = \sqrt{\frac{1}{2^2}\left[\left[10.0\times 0.00232\right]^2 + \left[10.0\times 0.00193\right]^2\right]} = 0.015\ \mu g/m^3$$

where $C_{\mathrm{Ref},j}$ is the certified reference concentration for the j^th bracketing procedure (usually consistent, but not required to be). For such a small number of bracketing tests, the appropriate way to calculate reproducibility is the Bound on Bias approach

$$u_{1\sigma}\left(Y_{FR}{}_H^*\right)_{\text{reproducibility}} = \frac{Max(Y_{FR_H}) - Min(Y_{FR_H})}{\sqrt{12}} = \frac{10.0 \times 0.972 - 10.0 \times 0.970}{\sqrt{12}} = 0.006\,\mu\text{g/m}^3$$

Assuming a vendor prime certification uncertainty of 0.05 μg/m^3

$$u_{1\sigma}\left(Y_{FR}{}_H^*\right)_{\text{reference}} = \overline{R} \times u_{1\sigma}\left(X_{V}{}_H^*\right)_{\text{certification}} = 0.971 \times 0.05 = 0.049\,\mu\text{g/m}^3$$

the combined certification uncertainty is

$$u_{1\sigma}\left(Y_{FR}{}_H^*\right)_{combined} = \sqrt{(0.015)^2 + (0.049)^2 + (0.006)^2} = 0.05\,\mu\text{g/m}^3$$

and the expanded certification uncertainty (k=2) is

$$u_{1\sigma}\left(Y_{FR}{}_H^*\right)_{\exp anded} = k \times u_{1\sigma}\left(Y_{FR}{}_H^*\right)_{combined} = 2 \times 0.05\,\mu\text{g/m}^3 = 0.10\,\mu\text{g/m}^3$$

Calculating the % uncertainty

$$\frac{u_{1\sigma}\left(Y_{FR}{}_H^*\right)_{\exp anded}}{Y_{FR}{}_H^*} \times 100\% = \frac{0.10\,\mu\text{g/m}^3}{9.71\,\mu\text{g/m}^3} \times 100\% = 1.0\%$$

As shown, the expanded, combined uncertainty is well within the 5% acceptance criterion of Section 6.4 of the interim elemental Hg gas traceability protocol[1].

B.12 Multi-Bracketing Uncertainty (Statistical Reproducibility)

When a certification process involves multiple bracketing tests, as in the example of Table B-7, reproducibility can be calculated using the statistical approach of Section A.4.1. The nesting calculation of this approach requires retaining the s_2 parameter (from the repeatability calculation, essentially the standard deviation of the individual ratios) in the table.

Table B-6. Dual-Bracketing Example Uncertainty Components

	Bracketing Test Number (j)					
Parameter	**1**	**2**	**3**	**4**	**5**	**6**
# of Brackets, K	3	3	3	3	3	3
Avg. Ratio, R_j	0. 972	0.970	0.967	0.965	0.968	0.971
$u_{1\sigma}\left(\overline{R}_j\right)_{linearity}$	0	0	0	0	0	0
$u_{1\sigma}\left(\overline{R}_j\right)_{stability}$	0.00078	0.00082	0.00076	0.00084	0.00074	0.00086
$u_{1\sigma}\left(\overline{R}_j\right)_{repeatability}$	0.00219	0.00175	0.00211	0.00184	0.00202	0.00193
$u_{1\sigma}\left(\overline{R}_j\right)$	0.00232	0.00193	0.00224	0.00202	0.00215	0.00211
Level-2 error, s_2	0.00381	0.00303	0.00364	0.00319	0.00348	0.00334

The statistical reproducibility formula is

$$u_{1\sigma}\left(Y_{FR_H^*}\right)_{\text{reproducibility}} = \frac{1}{\sqrt{J}}\sqrt{Max\left[0,s_3^2 - \frac{1}{K}\times s_{2,C_{\text{Cand}}}^2\right]}$$

where s_3 is simply the standard deviation of the candidate concentrations calculated from each of the six bracketing tests.

$s_3 = 0.0250\ \mu\text{g/m}^3$

and $s_{2,C_{\text{Cand}}}$ is pooled from the s_2 values based on the number of brackets in each set (K).

$$s_{2,C_{cand}} = \sqrt{\frac{1}{\sum K_j - J}\times\sum_{j=1}^{J}\left[(K_j - 1)\times\left(C_{\text{Ref},j}\times s_{2,j}\right)^2\right]}$$

When K and C_{Ref} are consistent between bracketing procedures, $s_{2,C_{\text{Cand}}}$ reduces to the root mean squared s_2 times the certified reference concentration ($C_{\text{Ref}} = Xv_H^* = 10.0\ \mu\text{g/m}^3$).

$$s_{2,C_{cand}} = 10.0\times\sqrt{\begin{matrix}(0.00381)^2 + (0.00303)^2 + (0.00364)^2 + \\ (0.00319)^2 + (0.00348)^2 + (0.00334)^2\end{matrix}\Big/6} = 0.0343\ \mu\text{g/m}^3$$

$$u_{1\sigma}\left(Y_{FR_H^*}\right)_{\text{reproducibility}} = \frac{1}{\sqrt{6}}\sqrt{Max\left[0,(0.0250)^2 - \frac{1}{3}\times(0.0343)^2\right]} = 0.006\ \mu\text{g/m}^3$$

Similarly to the dual-bracketing example, the comparison uncertainty is calculated by

$$u_{1\sigma}\left(Y_{FR_H^*}\right)_{comparison} = \sqrt{\frac{1}{J^2}\sum_{j=1}^{J}\left[C_{\text{Ref},j}\times u_{1\sigma}\left(\overline{R}_j\right)\right]^2} = \sqrt{\frac{1}{6^2}\begin{bmatrix}[10.0\times0.00232]^2 + [10.0\times0.00193]^2 + \\ [10.0\times0.00224]^2 + [10.0\times0.00202]^2 + \\ [10.0\times0.00215]^2 + [10.0\times0.00211]^2\end{bmatrix}} = 0.009\ \mu\text{g/m}^3$$

Assuming a reference standard uncertainty of 0.05 $\mu\text{g/m}^3$

$$u_{1\sigma}\left(Y_{FR_H^*}\right)_{\text{Reference}} = \overline{R}\times u_{1\sigma}\left(Xv_H^*\right)_{\text{Certification}} = 0.968\times0.05 = 0.048\ \mu\text{g/m}^3$$

the combined certification uncertainty is

$$u_{1\sigma}\left(Y_{FR_H^*}\right)_{combined} = \sqrt{(0.009)^2 + (0.048)^2 + (0.006)^2} = 0.05\ \mu\text{g/m}^3$$

and the expanded certification uncertainty (k=2) is

$$u_{1\sigma}\left(Y_{FR_H^*}\right)_{\exp anded} = k\times u_{1\sigma}\left(Y_{FR_H^*}\right)_{combined} = 2\times0.05\ \mu\text{g/m}^3 = 0.10\ \mu\text{g/m}^3$$

Calculating the % uncertainty

$$\frac{u_{1\sigma}\left(Y_{FR_H^*}\right)_{\exp anded}}{Y_{FR_H^*}} \times 100\% = \frac{0.10\,\mu\text{g/m}^3}{9.68\,\mu\text{g/m}^3} \times 100\% = 1.0\%$$

As shown, the expanded, combined uncertainty is well within the 5% acceptance criterion of Section 6.4 of the interim elemental Hg gas traceability protocol[1].

References

[1] "Interim EPA Traceability Protocol for Qualification and Certification of Elemental Mercury Gas Generators"; U.S. EPA; July 2009.

[2] *NIST/SEMATECH e-Handbook of Statistical Methods*, http://www.itl.nist.gov/div898/handbook, sec. 2.4.3.3 (2006).

[3] Levenson, M.S., Banks, D.L., Eberhardt, K.R., et al, "An Approach to Combining Results From Multiple Methods Motivated by the ISA GUM," J. Res. Natl. Inst. Stand. Technol., vol. 105, pp. 571-579 (2000).