Validation report of the data analysis template for the calibration of elemental mercury gas generators

Authors: Federica Gugole

Version: 1.0

Date: 14 September 2023

Introduction

As part of WP1 of the 19NRM03 SI-Hg project [1] Python3-powered Excel sheets have been developed. These are templates to be used for the data processing of the calibration data of mercury gas generators according to the *Protocol for the SI-traceable calibration of elemental mercury gas generators used in the field* [2]. Two templates have been developed: one for the single-point and one for the multi-point calibration. Both templates are suited for a setup with two channels (i.e., channel A and channel B). They should thus be modified in case of a different lab setup.

The single-point analysis reflects the steps indicated by NIST in their calibration protocol *Interim EPA Traceability Protocol for Qualification and Certification of Elemental Mercury Gas Generators* [3]. The multi-point analysis performs the single-point analysis for each setpoint and then perform a weighted least square regression to determine the calibration curve.

I report here the results of the developed single-point template when applied to the (single bracket) example reported in Section B of the *Companion Document* [4] in support of [3]. Additionally I report the results of a Weighted Least Squares (WLS) regression for a polynomial of order one performed both with the Python template and with R.

Validation of single-point analysis

Data setup

For the validation of the single-point template I used the data reported in Section B of the *Companion Document* in Table B-1. Since the template assumes that there are two channels in the setup, I used the same data both for channel A and for channel B. Differently from NIST protocol, our template assumes that there are no measurement repetitions in the data, thus the data sequence is of the type Ref-Cand-Ref-Cand-Ref and not Ref-Ref-Ref-Cand-Cand-Cand-Ref-Ref-Ref-Cand-Cand-Cand-Ref-Ref-Ref. So for the validation I used the Measurement Average reported in Table B-1 and not Readings #1-5. Furthermore I extended the Time reported in Table B-1 to a datetime format. The time passed between the measurements matches the Time reported in Table B-1. The measurements for channel B are added as if they happened in between the measurements of channel A. However, since channel A and channel B are processed separately, this does not affect the result. See Figure 1 for a snapshot of how the used dataset look like.

DateTime	Pk Area	Name	Valid
01/03/2023 09:00	0	Zero A	1
01/03/2023 09:05	0	Zero B	1
01/03/2023 09:10	10	Ref A	1
01/03/2023 09:15	10	Ref B	1
01/03/2023 09:20	9,85	Cand A	1
01/03/2023 09:25	9,85	Cand B	1
01/03/2023 09:30	10,2	Ref A	1
01/03/2023 09:35	10,2	Ref B	1
01/03/2023 09:40	9,96	Cand A	1
01/03/2023 09:45	9,96	Cand B	1
01/03/2023 09:50	10,25	Ref A	1
01/03/2023 09:55	10,25	Ref B	1
01/03/2023 10:00	10	Cand A	1
01/03/2023 10:05	10	Cand B	1
01/03/2023 10:10	10,4	Ref A	1
01/03/2023 10:15	10,4	Ref B	1
01/03/2023 10:20	0,32	Zero A	1
01/03/2023 10:25	0,32	Zero B	1

Figure 1: dataset used for the validation.

Input from user

Input information such as the reference expanded uncertainty, the reference concentration and the reproducibility uncertainty have been set to match the values used in the analysis reported in *Companion Document*. See Figure 2 for the details.

	Data type	Unit	
Reference concentration	float	$\mu g\ m^{-3}$	10
Extended uncertainty reference generator (k=2)	float	%	1,2
Reference generator channel A	string	-	Ref A
Reference generator channel B	string	-	Ref B
Candidate generator channel A	string	-	Cand A
Candidate generator channel B	string	-	Cand B
Zero channel A	string	-	Zero A
Zero channel B	string	-	Zero B
Apply zero correction	either 0 (False) or 1 (True)	-	1
Reproducibility uncertainty (k=1)	float	%	0,5
DateTime format	string	-	%d/%m/%Y %H:%M:%S

Figure 2: input information provided to perform the analysis.

Results

Since the same data have been used both for channel A and for channel B, I report here the results only for channel A (which are the same as the results for channel B). In Figure 3 an overview of the terms computed during the analysis is reported. Values are highlighted either in green if there is a match with the corresponding value reported in *Companion Document* or in yellow if they are similar but not the same as those reported in *Companion Document*. It can be noticed that there is a small disagreement for the stability uncertainty component and all the other components that are computed using the stability uncertainty. In particular I could not recover the values reported by NIST in Table B-4 using the formulas reported in the document. This issue has been discussed with Iris de Krom and Adriaan van der Veen and it has been decided to trust our computations. In any case the disagreement is very small and does not significantly affect the combined standard uncertainty which agrees quite well (and matches if rounded to the second decimal digit) with the combined certification uncertainty reported in *Companion Document*.

	A	B	C	D	E	F
1		**Symbol**	**Unit**	**Bracket_1**	**All_brackets**	
2	**Mean output ratio**	R		0,972164986	0,972164986	
3	**RSD output ratio**		%	0,397541102		
4	**Stability uncertainty**	u_stability	ng/m3	0,002369025		
5	**Repeatability uncertainty**	u_repeatability	ng/m3	0,001762957		
6	**Comparison uncertainty**	u_comparison	ng/m3		0,029530152	
7	**Reproducibility uncertainty**	u_reproducibility	ng/m3		0,048608249	
8	**Reference uncertainty**	u_reference	ng/m3		0,058329899	
9	**Combined std uncertainty**	u(c)	ng/m3		0,081468822	
10	**Expanded uncertainty (k=2)**	U(c)	ng/m3		0,162937643	
11	**Relative expanded uncertainty**	U(c)	%		1,676028716	
12	**Candidate concentration**	c	ng/m3		9,721649865	
13						

Figure 3: results of the analysis for channel A.

Validation of multi-point analysis

Data setup

Here we used the data used in the *multi_point_dummy.xlsm*. Since the procedure to obtain the candidate concentration and associated uncertainty for each setpoint is the same as in the single-point template, I assume that part to be validated. I thus take the output of that part as input for the validation of the WLS procedure. In particular, I note that due to rounding off happening when writing from Python to Excel, I had to get the values used in the regression by printing them to txt file (*input_data.txt*).

Results

We validated the results of the Weighted Least Squares Python routine *WLS* (part of the *statsmodels* library) by performing WLS regression in R (see script *validate_multipoint_regression.R*). We considered the case of a polynomial of order 1 (i.e., $y = a + bx$) and checked the values used for the polynomial selection (i.e., the corrected Akaike Information Criterion (AICc)) as well as the relevant statistics associated to the polynomial (i.e., parameters value and covariance, standard error, p-values and sum of the squared residuals). As it can be seen from Figures 4, 5 and 6, the values are nearly identical and the only differences are due to rounding off at a different decimal point when printing some values (which disappear if we round the respective numbers to the same decimal point).

Channel A	Parametei	Standard (	P-values
b_0	11,16746	5,803214	0,126635
b_1	0,932088	0,003763	1,59E-09

```
> print(df_chA)
  Parameters Standard error      P-values Sum squared residuals
1 11.1674638    5.803214008 1.266346e-01             0.5970865
2  0.9320876    0.003763302 1.594240e-09             0.5970865
>
```

Figure 6: comparison between the regression results obtained with Python (left) and with R (right).

Channel A	AICc	Sum squared residuals
poly0	95,25216	9157,581
poly1	42,42397	0,597086
poly2	48,36952	0,303785
poly3	76,83378	0,235183

```
> AICc(wls_chA)
[1] 42.42397
>
```

Figure 5: comparison between the AICc values computed with Python (left) and with R (right).

Covariance matrix		b_0	b_1
	b_0	33,67729	-0,02075
	b_1	-0,02075	1,42E-05

```
> vcov(wls_chA)
            (Intercept)       setpoint
(Intercept) 33.67729282 -2.075344e-02
setpoint    -0.02075344  1.416244e-05
>
```

Figure 4: comparison between the parameters' covariance computed with Python (left) and computed with R (right).

Conclusions

The single-point calibration template developed as part of the 19NRM03 SI-Hg project has been validated using sample data and related computations reported in *Companion Documentation* [4]. Overall the results match very well except for a small difference in the stability uncertainty (and thus the other quantities that use the stability uncertainty in their computation). This difference originates from the inability to reproduce the results reported in Table B-4 using the formulas reported in *Companion Documentation*. Since the difference is very small and does not significantly affect the end result, it has been decided that the current version of the template is acceptable and may be used for calibration services.

The Weighted Least Squares regression performed in the multi-point template as been validated against the WLS regression performed in R (given the same input). The results match almost perfectly thus concluding that the multi-point template is also validated.

Acknowledgments

This project 19NRM03 SI-Hg has received funding from the EMPIR programme co-financed by the Participating States and from the European Union's Horizon 2020 research and innovation programme.

References

[1] European Metrology Programme for Innovation and Research (EMPIR) project ‘Metrology for traceable protocols for elemental and oxidized mercury concentrations’ (19NRM03 SI-Hg).

[2] I. de Krom, A. van der Veen, F. Gugole, W. Corns, C. Roellig, T. Rajamäki, R. Moesler: Protocol for the SI-traceable calibration of elemental mercury gas generators used in the field

[3] Interim EPA traceability protocol for qualification and certification of elemental mercury gas generators, July 2009, NIST

[4] Companion Document Treatment of elemental mercury gas generator certification data and calculation of uncertainty, in support of [3], July 2009, NIST