



D2: Validation report for the calibration of elemental mercury gas generators including information on repeatability, reproducibility and uncertainty evaluation at emission and ambient levels extended to the sub  $\text{ng m}^{-3}$  level

---

*19NRM03 SI-Hg D2*

---

*Iris de Krom,<sup>1</sup> Adriaan van der Veen,<sup>1</sup> Federica Gugole,<sup>1</sup> Ruben Ziel,<sup>1</sup> Warren Corns<sup>2</sup>, Carsten Röllig<sup>3</sup>, Stefan Simon<sup>3</sup>*

<sup>1</sup> VSL, Department of Chemistry, Thijsseweg 11, 2629 JA Delft, the Netherlands

<sup>2</sup> PS Analytical Ltd, Arthur House, Crayfields Industrial Estate, Main Road, Orpington, Kent BR5 3HP, UK

<sup>3</sup> TÜV Rheinland, AM Grauen Stein, 51105 Köln, Germany

---

Lead partner: VSL

Version No.: 2

Due date deliverable: 31-07-2021

Actual submission date: 29-11-2023

This project 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.



The EMPIR initiative is co-funded by the European Union's Horizon 2020 research and innovation programme and the EMPIR Participating States

## Summary

In this deliverable the first validation results of the SI-Hg calibration protocol are reported. Within the SI-Hg project a protocol for the metrological calibration of elemental mercury gas generators used in the field was developed. For the validation the output of two different mercury gas generators was calibrated according to the protocol. As metrological reference standard the primary mercury gas standard from the Van Swinden Laboratory (VSL) was used. The measurements described in the protocol could be performed during the validation and the data was processed using a script to determine the output of the candidate generator and the uncertainty of the mercury concentration. Based on the validation measurements and data processing several improvements for the calibration protocol were identified and were used to improve the calibration protocol.

## Table of Contents

Summary .....	3
Table of Contents .....	4
1. Introduction .....	5
2. Equipment.....	7
2.1 Analyser.....	7
2.2 Zero air generator .....	7
2.3 Primary mercury gas generator .....	7
2.4 PSA 10.532 fixed output elemental mercury gas generator.....	7
2.4.1 Before modification .....	8
2.4.2 After modification.....	8
3. Measurements and methods.....	9
3.1 Measurements .....	9
3.2 Data processing.....	9
3.2.1 Precision: Repeatability and reproducibility.....	9
3.2.2 Data processing script.....	10
3.2.3 Output candidate generator .....	10
3.2.4 Interpolation function.....	10
3.2.5 Deviation .....	11
3.2.6 Uncertainty .....	11
4. Results and discussion .....	12
4.1 Measurement sequence .....	12
4.2 Multipoint comparison .....	14
4.2.1 Interpolation function.....	16
4.3 Single point comparison .....	18
4.3.1 Measurements April 2022.....	18
4.3.2 Measurements March 2023.....	22
4.3.3 Measurements June 2023.....	24
4.3.4 Measurements September 2023 .....	25
5. Conclusion.....	27
6. Improvements for the protocol .....	28
References .....	29

## 1. Introduction

Currently, mercury gas generators used in the field are not certified against primary standards and therefore lack traceability. The aim of the 19NRM03 SI-Hg project is to develop and validate metrological traceable protocols for the calibration of mercury gas generators used in the field (hereafter referred to as calibration protocol). In this way, the project will achieve significant improvements in the measurement comparability and uncertainty of mercury measurement results.

Although great efforts have been made in developing primary mercury standards and SI-traceable calibration methods for different mercury species, there are no standardised procedures that ensures the dissemination and uptake of the developed metrological traceability by calibration and testing laboratories and in the field. Scientifically sound calibration protocols, to determine the output of elemental mercury ( $\text{Hg}^0$ ) in the form of formally accepted documentary standards, are of fundamental importance to guarantee the accuracy and comparability of the mercury measurement data in Europe and globally. Furthermore, mercury gas generators certified using SI-traceable standards will provide the traceability and uncertainty needed by calibration and testing laboratories under ISO/IEC 17025:2017 accreditation to demonstrate their conformity in assessments.

$\text{Hg}^0$  mercury generators are used for the calibration and zero and span checks of  $\text{Hg}^{\text{tot}}$  analysis in the field. However, the output of the mercury generators does not have metrological traceability in Europe. To guarantee the quality assurance of mercury measurements in the field, in the USA, the United States Environmental Protection Agency (US EPA) in collaboration with the National Institute of Standards and Technology (NIST) have developed interim protocols for establishing traceability of  $\text{Hg}^0$  generators [1]. Taking the EPA protocol as starting point the SI-Hg project developed their own calibration protocol [2].

The SI-Hg calibration protocol specifies the procedures for establishing traceability to the SI units for the quantitative output of elemental mercury generators that are employed in regulatory applications for emission monitoring or testing.

This protocol provides methods for

- the experimental procedures to compare the output of elemental mercury gas generators
- the data processing for determination of mercury concentration and the expanded uncertainty of the mercury concentration obtained from the elemental mercury gas generator.

The mercury concentration in a gas mixture prepared with a mercury gas generator is determined by comparison with a metrologically traceable reference standard to calibrate the output of a candidate generator.

The comparison can be performed at one concentration level (single-point calibration) or at several concentration levels (multipoint calibration) using a bracketing sequence. When applying the bracketing sequence, the outputs from the generators are introduced alternately to a mercury analyser so that each response from the candidate generator is bracketed by a pair of responses from the reference generator. At each concentration level, the bracketing procedure requires a minimum of four responses from the certified reference standard generator and three responses from the candidate generator (Table 1).

Table 1 – An example of a bracketing procedure injection sequence for a given mercury concentration

Generator ID	Response ID
Reference standard	Ref <sub>1</sub>
Candidate	Cand <sub>1</sub>
Reference standard	Ref <sub>2</sub>
Candidate	Cand <sub>2</sub>
Reference standard	Ref <sub>3</sub>
Candidate	Cand <sub>3</sub>
Reference standard	Ref <sub>4</sub>

There are two approaches for the data processing:

- Without zero correction
- With zero correction

Zero correction is required only for an analyser that does not perform an automatic baseline correction for every reading.

The mercury concentrations are expressed in ng m<sup>-3</sup> or µg m<sup>-3</sup> (at 20 °C and 101.325 kPa).

In this deliverable the first validation results of the calibration protocol are reported. For the validation, Van Swinden Laboratory (VSL) used its primary mercury gas standard [3-5]. The standard consists of a generator of elemental mercury vapor and a set of thermal mass flow controllers able to achieve the concentration range required for the validation. The reference generation method is in compliance with ISO 6145-8:2005, the diffusion method [6]. The metrological traceability and the accuracy of the mercury concentrations is ensured. A generator from P S Analytical (PSA) was used as the candidate generator. The PSA generator was calibrated at the VSL against the primary mercury gas standard. Thereafter the calibrated generator was used at Technischer Überwachungsverein Rheinland Energy GmbH (TÜV) to calibrate a second PSA generator. As a last step the PSA generator was calibrated again at VSL against the primary mercury gas standard.

During the validation the precision of the gas generator output was determined. Furthermore, a script was developed to [7]:

- calculate the output mercury concentration of the candidate generator
- determine the interpolation function for the output mercury concentration in case of a multipoint calibration
- calculate the deviation between the candidate generator setpoint and the calculated output mercury concentration
- calculate the uncertainty of the calculated candidate generator output mercury concentration

The measurement results reported here are performed in the µg m<sup>-3</sup> range. Measurements in the ng m<sup>-3</sup> range were reported in D5 of the SI-Hg project [8] and reported in literature [9].

## 2. Equipment

### 2.1 Analyser

The output of the reference generator and the candidate generator are analysed with a PSA Sir Galahad II (PSA SG) analyser. The analyser is designed for measuring mercury in gaseous samples. To monitor continuously for mercury in air or gas, two Amasil® traps are employed in parallel. While sample gas is pulled over one trap, to absorb any mercury present, the second trap is analysed. Mercury is measured using atomic fluorescence detection. The analyser is accompanied by a stream selector which can be used to connect 4 different mercury gas generators for alternately measurement. By selecting different GAIN levels (1, 10 or 100) the whole range mercury concentrations  $0.1 \text{ ng m}^{-3} - 100 \text{ } \mu\text{g m}^{-3}$  can be analysed. Measurement uncertainty of the Sir Galahad analyser is 2 % [10].

As the analyser has two traps (“A” and “B”) to concentrate the sample, the results obtained with each trap (A or B) are processed separately.

### 2.2 Zero air generator

For the generation of zero air the Tekran Model 1100 Zero Air Generator is used. Ambient air is drawn into the generator through a  $0.1 \text{ } \mu\text{m}$  particulate filter by a pump, and sent to a multi-component mercury scrubber. The cleaned air is then passed through another  $0.1 \text{ } \mu\text{m}$  filter before being sent to the outlet. Via the stream selector the zero air is directed to the analyser to obtain zero responses ( $r_{zero}$ ) during the bracketing measurement sequence described in the calibration protocol.

### 2.3 Primary mercury gas generator

The primary gas standard was developed as an elemental mercury ( $\text{Hg}^0$ ) gas generator to establish metrological traceability of mercury concentration measurement results, based on a gravimetric approach, for ambient air levels as well as higher concentrations [3-5].

The working principle of the primary mercury vapour generator is based on diffusion according to ISO 6145-8:2005 [4, 6]. This is a gravimetric method to provide traceability to the International System of Units (SI) for concentration measurement results of mercury. Using specially designed diffusion cells,  $\text{Hg}^0$  is evaporated under well-controlled conditions (temperature, flow rate and pressure). By weighing the diffusion cells at regular time intervals with a high-resolution balance, an accurate mercury diffusion flow rate is obtained.

In the generator, the diffusion cells are housed in a diffusion chamber. The diffusion chamber is temperature ( $20.0 \text{ } ^\circ\text{C} \pm 0.1 \text{ } ^\circ\text{C}$ ) and pressure ( $105.0 \text{ kPa} \pm 0.1 \text{ kPa}$ ) controlled. At the bottom, a nitrogen flow of  $500 \text{ mL min}^{-1}$  enters the diffusion chamber. All of the flow, enriched by mercury vapour, is then guided to the outlet of the diffusion chamber through an aperture at the top. Standard  $\text{Hg}^0$  gas mixture concentrations are prepared by mixing the  $\text{Hg}^0$  vapours in nitrogen in a second step with flows, between  $1 \text{ L min}^{-1}$  and  $25 \text{ L min}^{-1}$ , of complementary gas, e.g., purified air.

### 2.4 PSA 10.532 fixed output elemental mercury gas generator

P S Analytical (PSA, United Kingdom) provided the 10.523 elemental Hg generator which was used for the validation of the calibration protocol at standard conditions of temperature (273.15 K and pressure (101.325 kPa).

The generator includes a mercury reservoir. The unit operates on the principle of dilution a saturated source of mercury at a known temperature according to ISO 6145-9 [11]. A fixed flow rate is passed across the mercury reservoir ensuring that the gas becomes saturated with mercury. The

mercury-saturated gas is then diluted with a second fixed flow rate to obtain a single mercury concentration.

#### 2.4.1 Before modification

Initial tests were performed using a reservoir critical orifice operated at 3psig (20.68 kPa) inlet pressure. This arrangement was highly dependent on the backpressure applied to the output of the generator. It was therefore modified to use a critical orifice which operated at higher pressure. The setpoint of the generator using this arrangement determined by PSA was 11876 ng m<sup>-3</sup>.

#### 2.4.2 After modification

The final tests were performed using a critical orifice operated at 30psig (206.8 kPa). This arrangement was relatively independent to the backpressure applied to the output of the generator as the pressure differential across the orifice was much higher than the backpressure. The setpoint of the generator using this arrangement determined by PSA was 10640 ng m<sup>-3</sup>.



## 3. Measurements and methods

### 3.1 Measurements

During the validation comparisons, different mercury concentrations were tested (Table 2) as described in calibration protocol [2].

Table 2 – Comparisons performed for the validation of the calibration protocol including the reference generator mercury concentration(s) used ( $c_{ref}$ ) and the setpoint(s) of the candidate generator ( $c_{cand}$ ).

Comparison	Reference generator	$c_{ref}$ range (ng m <sup>-3</sup> )	$U(c_{ref})$ (%) ( $k = 2$ )	Candidate generator	$c_{cand}$ (ng m <sup>-3</sup> )	$U(c_{cand})$ (%) ( $k = 2$ )	Partner	Date
Multipoint	VSL 3 x Ø3 mm diffusion cells	900 – 2500	5	VSL 1 x Ø8 mm diffusion cell	900 – 2500	5	VSL	July 2022
Single point	VSL 1 x Ø33 mm diffusion cell	4000 - 12000	4	PSA 10.532 before modification	11876	4.9	VSL	April 2022
Single point	VSL 1 x Ø33 mm diffusion cell	8000 - 14000	4	PSA 10.532 after modification	10640	4.9	VSL	March 2023
Single point	PSA 10.532 after modification VSL calibrated output	11393	4	PSA multipoint 10.536	11350	4.8	TUV	July 2023
Single point	VSL 1 x Ø33 mm diffusion cell	8000 - 14000	4	PSA 10.532 after modification (final tests)	10640	4.9	VSL	September 2023

For the multipoint comparison the VSL generator with the validated 3 x Ø3 mm diffusion cells was used as reference generator and the VSL generator with 1 new, not validated, Ø8 mm diffusion cell was used as candidate generator. With both the reference and the candidate 6 mercury concentrations were generated comparable within 10 %. This comparison was repeated on 3 different days.

For the single point comparison the VSL generator was used as reference generator and the PSA 10.532 fixed output generator as the candidate. With the reference generator several different mercury concentrations were obtained and the fixed output of the candidate generator was compared with the different mercury concentrations from the reference generator. This comparison was repeated on 3 different days. After modification of the generator the measurements were repeated using nitrogen and air as complementary gas. At TUV the PSA generator was used as the reference generator and a candidate generator, a PSA multipoint 10.536 generator, was calibrated at a single point. In this way a calibration chain was obtained. For a last check the PSA fixed output generator was calibrated one last time at VSL against the VSL reference standard.

### 3.2 Data processing

#### 3.2.1 Precision: Repeatability and reproducibility

To determine the precision of the output of the candidate generator the repeatability and reproducibility standard deviations was determined.

The repeatability of the output is the closeness of the agreement between the results of successive individual measurements of Hg<sup>0</sup> concentrations generated by the candidate generator carried out under the same conditions of measurement.

The reproducibility of the output is the closeness of the agreement between the results obtained on 4 different days in a one-month period.

Data obtained during calibration of the candidate generator will be used to determine the repeatability and reproducibility standard deviations.

The repeatability standard deviation ( $s_r$ , expressed as coefficient of variation in %) and within-laboratory reproducibility standard deviation ( $s_R$ , expressed as coefficient of variation in %) was calculated according to ISO 5725-2:2019 using one-way analysis of variance (ANOVA) [12].

### 3.2.2 Data processing script

For the processing of the data a script was developed using Python 3.8.8 [7]. The script was used to calculate different parameters based on the responses obtained with during the bracketing measurement sequence:

- Output mercury concentration(s) of the candidate generator ( $c$ ) at the setpoint(s) ( $c_{cand}$ )
- In case of a multi-point calibration: the interpolation function for the output mercury concentration of the candidate generator as a function of the setpoint
- Deviation between the candidate generator setpoint and the calculated output mercury concentration
- Uncertainty of the calibrated candidate generator output mercury concentration ( $U(c)$ )

The software can be used to process the data with and without zero correction.

### 3.2.3 Output candidate generator

Based on the analyser responses obtained from the bracketing measurement sequence the reference generator output mercury concentration ( $c$ ) at the setpoint can be calculated as described in the calibration protocol (see 8.6.1, [2]). The data can be processed with zero correction and without zero correction. An analyser without automatic baseline corrections was used, therefore all results were calculated using zero correction.

### 3.2.4 Interpolation function

When a candidate gas generator was calibrated at multiple points, the generator can be used 1) only at the calibrated set points or 2) an interpolation function can be used to calculate the output of the generator at any given set point in the range the generator was calibrated. For option 2 an interpolation function needs to be determined based on the calibration data as described in the calibration protocol (see Annex 2, [2]). Different types of interpolation functions can be selected (Table 3).

Table 3 – Interpolation functions

Polynomial	Degree	Name	Function
Poly 0	0	Non-zero constant	$c = c_{cand}$
Poly 1	1	Linear	$c = b_0 + b_1 c_{cand}$
Poly 2	2	Quadratic	$c = b_0 + b_1 c_{cand} + b_2 c_{cand}^2$
Poly 3	3	Cubic	$c = b_0 + b_1 c_{cand} + b_2 c_{cand}^2 + b_3 c_{cand}^3$

To determine which function is the best fit for the calibration data statistical test can be used as explained in Annex 2 of the calibration protocol [2].

Based on the statistical tests the interpolation function best fit for the calibration data can be selected.

### 3.2.5 Deviation

The deviation ( $D_{xi}$ ,  $\mu\text{g m}^{-3}$ ) and relative deviation ( $D_{rel}$ , %) is the closeness of the candidate generator setpoint ( $c_{cand}$ ) from the calculated candidate generator output mercury concentration ( $c$ ). For all the measurement results obtained the deviation was determined (Equations (1) and (2)).

$$D_c = c - c_{cand} \quad \text{eq. (1)}$$

$$D_{rel} = \frac{D_c}{c_{cand}} \quad \text{eq. (2)}$$

### 3.2.6 Uncertainty

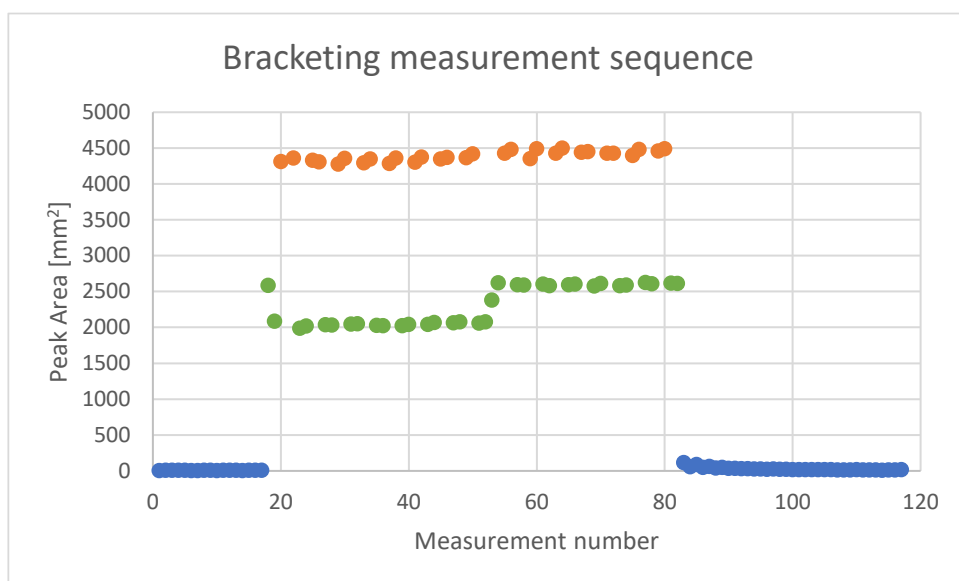
The uncertainty of the calculated candidate generator output mercury concentration ( $U(c)$ ) is calculated as described in the calibration protocol (see Annex 1 [2]) using the software developed for the data processing (see 3.2.3).

## 4. Results and discussion

### 4.1 Measurement sequence

According to the calibration protocol the comparison is performed using the bracketing sequence [2]. When applying the bracketing sequence, the outputs from the generators are introduced alternately to an analyser so that each response from the candidate generator is bracketed by a pair of responses from the reference standard generator. At each concentration level, the bracketing procedure requires a minimum of four responses, on each channel in case of a continuous sampling analyser, from the certified reference standard generator and three responses from the candidate generator. The bracketing sequence also starts and ends with introducing zero gas to the analyser. The response from the zero gas is used during the data processing with zero correction. During the data processing for each concentration level calculate the relative standard deviation (RSD) of the output ratios (R). The RSD shall not exceed 2.0%. If the RSD value is exceeded, the test is invalid and shall be repeated.

Below an example of the general trend observed for all measurements. When starting up the analyser the zero measurements give a stable response (Figure 1 top). If the response is stable only 1 measurement, on each channel in case of a continuous sampling analyser, is needed. The first measurements of the reference generator and the candidate generator show instable results during the validation measurements (Figure 1 bottom). In this case these measurements were discarded until a stable signal is obtained otherwise the RSD of the R value will be  $> 2.0\%$ .



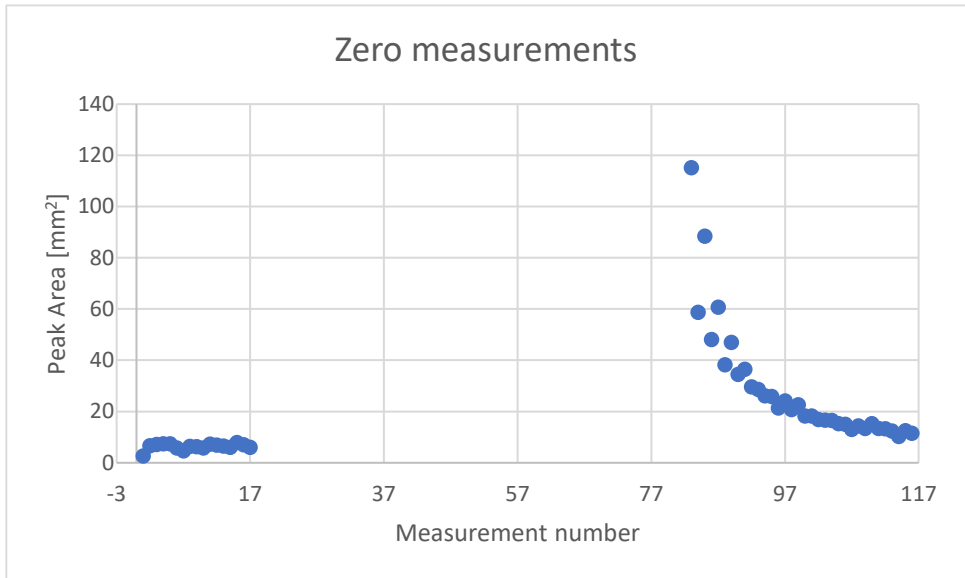


Figure 1: Plots of the measurement sequence (top), zoom in of the zero measurements (bottom). Blue dots zero measurements, green dots reference generator measurements, orange dots candidate generator measurements.

For the data processing of the data in Figure 1 the first measurements, on each channel, are regarded as instable and were discarded. Based on the other measurement the R values and the RSD of the R values was calculated according to the calibration protocol [2]. After bracket number 2 the RSD is calculated over 2 R values, after bracket number 3 over 3 R values and so on (Table 4).

Table 4 – Overview of R values obtained and the corresponding RSD value for Channel A measurements in Figure 1.

Bracket number	R	RSD (%)	R	RSD (%)
1	2.154		1.677	
2	2.098	1.9	1.706	1.2
3	2.112	1.4	1.721	1.3
4	2.120	1.1	1.719	1.2
5	2.119	1.0	1.692	1.1
6	2.121	0.9	1.703	1.0
7	2.123	0.8	1.718	1.0

When more R values are used the RSD value decreases. Which will eventually lead to a smaller contribution in the uncertainty of  $c$ , (see Annex 1 in the calibration protocol [2]). Therefore, as a minimum triplicate responses for the mercury concentration from the candidate generator and quadruple responses for the mercury concentration from the reference standard must be obtained after the signal is stable. It is advised to perform more measurements to reduce the RSD.

Furthermore, in this example (Figure 1), at the end of the sequence 15 measurements are needed, on each channel, before the zero measurements are stable. Closing the measurement sequence with one zero measurement is not enough. The zero measurements need to be continued until a stable value is obtained. Once the measurements are stable a single measurement result, for each channel, can be used for the data processing. The zero measurements at the start of the sequence usually show stable results immediately. Nevertheless, it is advised to perform several measurements to ensure a stable signal. To calculate a representative R value for the zero measurements the median of

the zero measurements is used for the calculations in further data processing using the software developed.

#### 4.2 Multipoint comparison

During this comparison the VSL generator was used as the reference standard as well as the candidate generator. The VSL generator consist of two systems which can be used at the same time. For the reference generator 3 x Ø3 mm diffusion cells are used to generate mercury concentrations. These cells were fully characterised [6]. For generation of mercury concentrations with the candidate generator 1 x Ø8 mm diffusion cell was used, this cell was not fully characterised and the mercury diffusion rate is therefore an estimation. With both the reference standard and the candidate generator 5 mercury concentrations were generated comparable within 10 % (Table 2). This measurement sequence was repeated 3 times (day 1, 2 and 3) in a period of 6 months.

To calculate the results the first responses obtained with the bracketing sequence were discarded. For the calculations at least triplicate responses for the mercury concentration from the candidate generator and quadruple responses for the mercury concentration from the reference standard were used. The responses were used to calculate the actual mercury concentration and the uncertainty according to the calibration protocol [2]. Zero correction was applied. The raw data and the results obtained can be found in a repository [13].

The RSD values of the R values are all  $\leq 2.0$  % and have an average value of 0.4 % and a maximum of 1.5 % for both channel A and channel B.

The mercury concentration,  $c$ , obtained with the candidate generator, the uncertainty,  $U(c)$ , of the concentration and the deviation,  $D_c$  and  $D_{rel}$ , compared to the setpoint were calculated according to the procedure in the calibration protocol using the data processing software (Table 5, Figure 2).

Table 5 – Summary of the results obtained for the mercury concentration,  $c$ , from the candidate generator, the uncertainty,  $U(c)$ , of the concentration and the deviation,  $D_c$  and  $D_{rel}$ , compared to the setpoint for channel A and channel B.

Measurement day	$c_{ref} \pm U$	$c_{cand}$	$c$	$U(c) (k = 2)$		Deviation	
	(ng m <sup>-3</sup> ) ( $k = 2$ )	(ng m <sup>-3</sup> )	(ng m <sup>-3</sup> )	(ng m <sup>-3</sup> )	(%)	(ng m <sup>-3</sup> )	(%)
<b>Channel A</b>							
1	975 ± 49	1071	1013	52	5.1	-58	-5.6
	975 ± 49	1367	1283	65	5.1	-84	-6.3
	1271 ± 64	1675	1578	84	5.3	-97	-6.0
	1578 ± 79	1964	1841	95	5.2	-123	-6.5
	1868 ± 93	2263	2131	115	5.4	-132	-6.1
	2167 ± 108	2563	2411	122	5.1	-152	-6.2
2	2226 ± 111	1071	993	50	5.0	-78	-7.3
	979 ± 49	1367	1270	67	5.2	-97	-7.1
	1282 ± 64	1675	1546	80	5.2	-129	-7.7
	1601 ± 80	1964	1813	92	5.1	-151	-7.7
	1907 ± 95	2263	2096	113	5.4	-167	-7.4
	2226 ± 111	2563	2378	127	5.3	-185	-7.2
3	2226 ± 111	1071	1015	51	5.1	-56	-5.2
	979 ± 49	1367	1283	70	5.4	-84	-6.1
	1282 ± 64	1675	1539	81	5.3	-136	-8.1
	1601 ± 80	1964	1801	92	5.1	-163	-8.3
	1907 ± 95	2263	2107	109	5.2	-156	-6.9

	2226 ± 111	2563	2368	126	5.3	-195	-7.6
Measurement day	$c_{ref} \pm U$	$c_{cand}$	$c$	$U(c) (k = 2)$		Deviation	
	(ng m <sup>-3</sup> ) (k = 2)	(ng m <sup>-3</sup> )	(ng m <sup>-3</sup> )	(ng m <sup>-3</sup> )	(%)	(ng m <sup>-3</sup> )	(%)
<b>Channel B</b>							
1	975 ± 49	1071	1011	52	5.1	-60	-5.6
	975 ± 49	1367	1286	66	5.2	-81	-6.3
	1271 ± 64	1675	1564	81	5.2	-111	-6.0
	1578 ± 79	1964	1843	95	5.1	-121	-6.5
	1868 ± 93	2263	2128	110	5.2	-135	-6.1
	2167 ± 108	2563	2421	129	5.3	-142	-6.2
2	2226 ± 111	1071	987	50	5.0	-84	-7.3
	979 ± 49	1367	1270	65	5.1	-97	-7.1
	1282 ± 64	1675	1554	79	5.1	-121	-7.7
	1601 ± 80	1964	1807	91	5.1	-157	-7.7
	1907 ± 95	2263	2128	110	5.1	-135	-7.4
	2226 ± 111	2563	2403	127	5.3	-160	-7.2
3	2226 ± 111	1071	1041	53	5.1	-30	-5.2
	979 ± 49	1367	1272	65	5.1	-95	-6.1
	1282 ± 64	1675	1552	89	5.8	-123	-8.1
	1601 ± 80	1964	1833	94	5.1	-131	-8.3
	1907 ± 95	2263	2095	107	5.1	-168	-6.9
	2226 ± 111	2563	2426	127	5.2	-137	-7.6

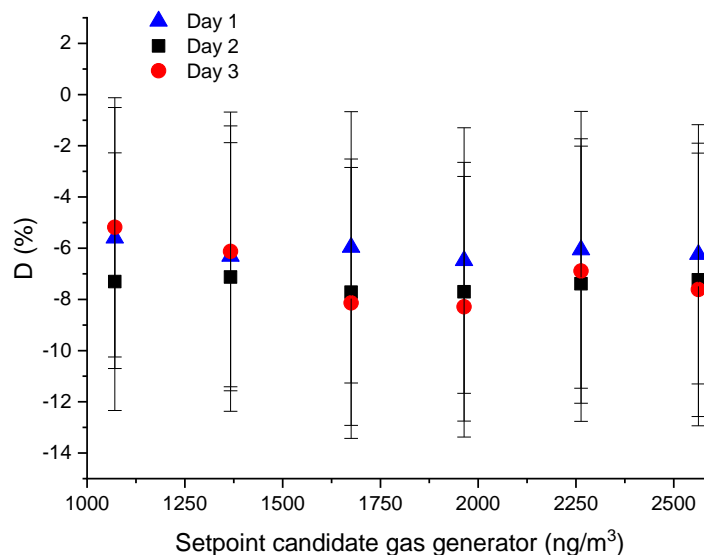


Figure 2 – Plot with the setpoints of the candidate generator against the deviation of the output of the candidate generator from the setpoint compared to the reference generated. Results from the three measurement days from Channel A are shown.

The uncertainties calculated ( $U(c)$ ) range from 5.0 to 5.8 %. The most important uncertainty source is the uncertainty of the reference standard of 5.0 %. The uncertainty increased in this case with a maximum of 0.8 % due to the stability and repeatability of the measurements or the comparison uncertainty. Another option is to add a contribution to the uncertainty for the reproducibility. For

the measurements performed during the validation and later we set the reproducibility uncertainty to zero. The results from the 3 measurement days and from Channel A and Channel B are comparable within the obtained uncertainty of 5% (Figure 2). Therefore, it was not necessary to add extra uncertainty due to the reproducibility of the measurements. When the measurements, especially those obtained on different days, are not comparable within the uncertainty it is advised to add uncertainty for the reproducibility to the uncertainty calculation.

The average deviation between the setpoint of the candidate generator ( $c_{cand}$ ) and the calculated concentration ( $c$ ) was -7 %. As the results are comparable the results of Channel A and Channel B and the 3 measurement days can be averaged for a calibration report.

#### 4.2.1 Interpolation function

For the data sets from each measurement day and each channel the interpolation function and regression coefficients were determined. Based on the weighted squared deviation and the Akaike Information Criterion for small sample size (AICc) the interpolation function with the best fit of the data was determined. Different interpolation functions or polynomials can be assigned (Table 3). As an example, the results of day 3 are discussed here (Table 6, Table 7 and Figure 3). The raw data and the results obtained can be found in a repository [13].

Table 6 – Results of the AICc test for the different polynomials.

	<b>Channel A</b>	<b>Channel B</b>
Poly 0	96	96
<u>Poly 1</u>	<b>55</b>	65
Poly 2	62	<b>63</b>
Poly 3	90	90

Table 7 – Results of the sum squared residuals

	<b>Channel A</b>	<b>Channel B</b>
Poly 0	15249	12993
<u>Poly 1</u>	<b>7</b>	32
Poly 2	4	<b>5</b>
Poly 3	3	3



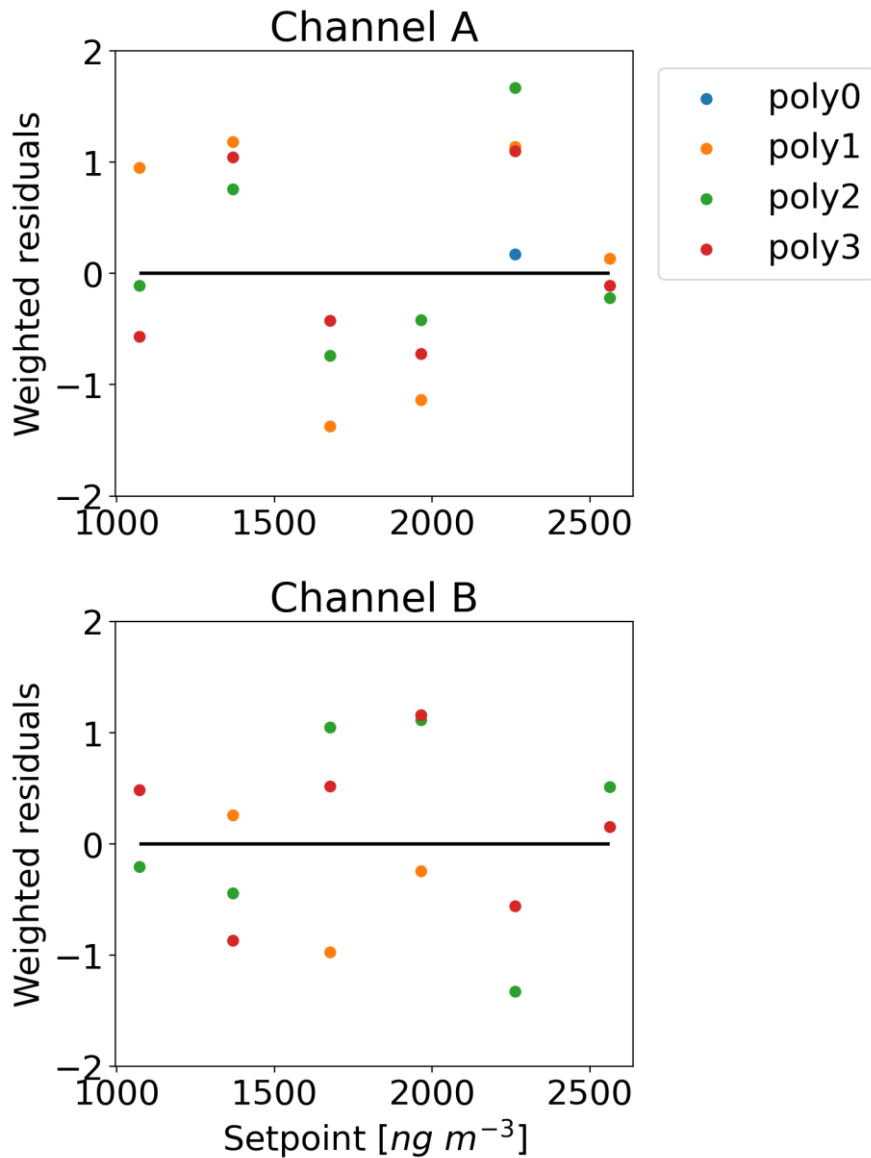


Figure 3 – Graphic representation of the weighted residuals compared to the setpoint of the candidate generator.

The polynomial with the smallest value in the AICc test is the best fit for the data. For Channel A, a linear function is the best fit and Channel B this is a quadratic function. Furthermore, a function is acceptable if the weighted residuals are in absolute value  $\leq 2$  (Figure 3). For Channel A poly 1, poly 2 and poly 3 the weighted residuals are in absolute value  $\leq 2$ . The best fit for the data is poly 1 as the data will be overfitted when using poly 2 or poly 3. Statistical tests give the same results for the best fit of the data (Table 8).

Table 8 – Regression coefficients for the interpolation function of Channel A and Channel B.

Channel A	Coefficients	Standard error
$b_0$	9.5	23
$b_1$	0.920	0.010
Channel B	Coefficients	Standard error
$b_0$	375	99
$b_1$	0.50	0.11
$b_2$	0.000116	0.000028

To determine if the regression coefficients are comparable for Channel A and Channel B the functions can be averaged and a value for the comparability can be calculated (Chi squared). When they are comparable they can be combined (Figure 4).

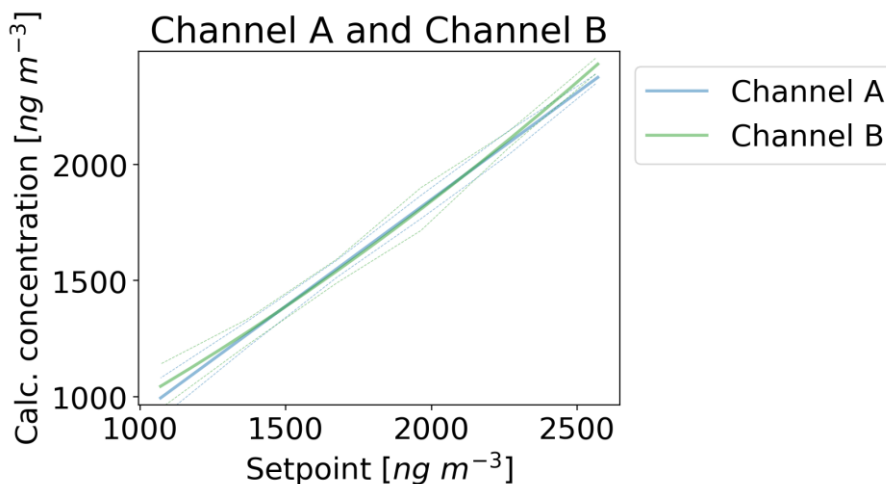


Figure 4 – Overlay interpolation functions channel A and channel B.

In this example the average regressed function is acceptable with a Chi squared value of 4.3 and the probability that such a chi squared value should occur by chance is 0.93. Even though the optimal function for Channel A and Channel B is different they can be combined to a linear function. The contribution of an extra parameter for the data from channel B is small as can be seen in the small decrease of the AICc (Table 6) and the sum squared residuals (Table 7).

The average regression coefficients for Channel A and Channel B and a covariance matrix were calculated (Table 9 and Table 10). The covariance matrix explains how the two data sets ( $b_0$ ,  $b_1$ ) are correlated. This correlation should be included in the calculation of the total uncertainty associated to the interpolation function [2].

Table 9 – Average regression coefficients

	Parameters	Standard error
$b_0$	-1.8	36
$b_1$	0.930	0.016

Table 10 – Covariance matrix

	$b_0$	$b_1$
$b_0$	1268	-0.547
$b_1$	-0.547	0.000246

### 4.3 Single point comparison

#### 4.3.1 Measurements April 2022

During this comparison the VSL generator was used as the reference generator and the PSA 10.532 fixed generator as the candidate generator. With the reference generator, 4 mercury concentrations were obtained and each concentration was compared with the output of the candidate generator,

which has a fixed mercury concentration output of 11876 ng m<sup>-3</sup> (Table 2). This sequence was repeated 4 times (day 1, 2, 3 and 4) in a period of 1 month. As the output of the generator was influenced by backpressure created by the sampling line from the generator to the analyser, the calibration was repeated with smaller tubing length. To study the influence of the complementary gas the calibration with shorter tubing has also been performed with air as complementary gas instead of nitrogen. During all experiments the backpressure was monitored with a GE Druck PACE 1000 pressure meter. No correction for the backpressure was applied as a relationship between the output of the generator and the backpressure is not known.

To calculate the results the first responses obtained with the bracketing sequence were discarded until a stable response was obtained. To determine the precision of the PSA generator the peak areas obtained during the 4 measurement days were used to calculate the repeatability standard deviation ( $s_r$ , expressed as coefficient of variation in %) and reproducibility standard deviation ( $s_R$ , expressed as coefficient of variation in %) according to ISO 5725-2:2019 using one-way analysis of variance (ANOVA) [12] (Table 11).

Table 11 – Results of  $s_r$  and  $s_R$  for the PSA generator based the peak areas recorded with Channel A and Channel B of the PSA SG analyser.

Channel	$s_r$ (%)	$s_R$ (%)
A	0.6	1.8
B	0.6	2.0

The PSA generator has a precision of <2 %, the uncertainty of the PSA SG analyser (2 %) is probably the most important source for the precision.

For calculation of the mercury concentration,  $c$ , at least triplicate responses for the mercury concentration from the candidate generator and quadruple responses for the mercury concentration from the reference standard were used. The responses were used to calculate the calibrated mercury concentration and the uncertainty according to the calibration protocol [2]. A zero correction was applied. The raw data and the results obtained can be found in a repository [13].

The RSD values of the R values are all  $\leq 2.0$  %. The mercury concentration,  $c$ , obtained with the PSA generator, the uncertainty,  $U(c)$ , of the concentration and the deviation,  $D_c$  and  $D_{rel}$ , compared to the setpoint were calculated according to the procedure in the calibration protocol using the data processing software (Table 12).

Table 12 – Summary of the results obtained for the mercury concentration,  $c$ , obtained with the PSA generator, the uncertainty,  $U(c)$ , of the concentration and the deviation,  $D_c$  and  $D_{rel}$ , compared to the setpoint of 11876 ng m<sup>-3</sup> for channel A and channel B.

Measurement day	$c_{ref} \pm U$	$c$	$U(c) (k = 2)$		Deviation	
	(ng m <sup>-3</sup> ) (k = 2)	(ng m <sup>-3</sup> )	(ng m <sup>-3</sup> )	(%)	(ng m <sup>-3</sup> )	(%)
<b>Channel A</b>						
1	4496 ± 180	9255	384	4.2	-2621	-22.1
	5642 ± 226	9237	374	4.0	-2639	-22.2
	7572 ± 303	9236	377	4.1	-2640	-22.2
	11509 ± 460	9503	386	4.1	-2373	-20.0
2	4496 ± 180	9575	387	4.0	-2301	-19.4
	5642 ± 226	9491	383	4.0	-2385	-20.1
	7572 ± 303	9656	394	4.1	-2220	-18.7

	11509 ± 460	9536	388	4.1	-2340	-19.7
3	4496 ± 180	9311	382	4.1	-2565	-21.6
	5642 ± 226	9751	392	4.0	-2125	-17.9
	7572 ± 303	9910	405	4.1	-1966	-16.6
	11509 ± 460	10043	438	4.4	-1833	-15.4
4	4496 ± 180	9541	382	4.0	-2335	-19.7
	5642 ± 226	9658	392	4.1	-2218	-18.7
	7572 ± 303	9884	403	4.1	-1992	-16.8
	11509 ± 460	9761	411	4.2	-2115	-17.8
Shorter tubing	11509 ± 460	10575	437	4.1	-2621	-22.1
Shorter tubing air	11509 ± 460	9804	457	4.7	-2639	-22.2
<b>Measurement day</b>	<b><math>c_{ref} \pm U</math></b>	<b><math>c</math></b>	<b><math>U(c) (k = 2)</math></b>		<b>Deviation</b>	
	( $\mu\text{g m}^{-3}$ ) ( $k = 2$ )	( $\text{ng m}^{-3}$ )	( $\text{ng m}^{-3}$ )	(%)	( $\text{ng m}^{-3}$ )	(%)
<b>Channel B</b>						
1	4496 ± 180	9243	377	4.1	-2633	-22.2
	5642 ± 226	9297	384	4.1	-2579	-21.7
	7572 ± 303	9255	378	4.1	-2621	-22.1
	11509 ± 460	9515	386	4.1	-2361	-19.9
2	4496 ± 180	9552	395	4.1	-2324	-19.6
	5642 ± 226	9521	385	4.0	-2355	-19.8
	7572 ± 303	9666	389	4.0	-2210	-18.6
	11509 ± 460	9510	393	4.1	-2366	-19.9
3	4496 ± 180	9516	405	4.3	-2360	-19.9
	5642 ± 226	9738	393	4.0	-2138	-18.0
	7572 ± 303	9837	408	4.1	-2039	-17.2
	11509 ± 460	10027	411	4.1	-1849	-15.6
4	4496 ± 180	9603	387	4.0	-2273	-19.1
	5642 ± 226	9651	396	4.1	-2225	-18.7
	7572 ± 303	9885	421	4.3	-1991	-16.8
	11509 ± 460	9635	398	4.1	-2241	-18.9
Shorter tubing	11509 ± 460	10475	422	4.0	-2633	-22.2
Shorter tubing air	11509 ± 460	9702	405	4.2	-2579	-21.7

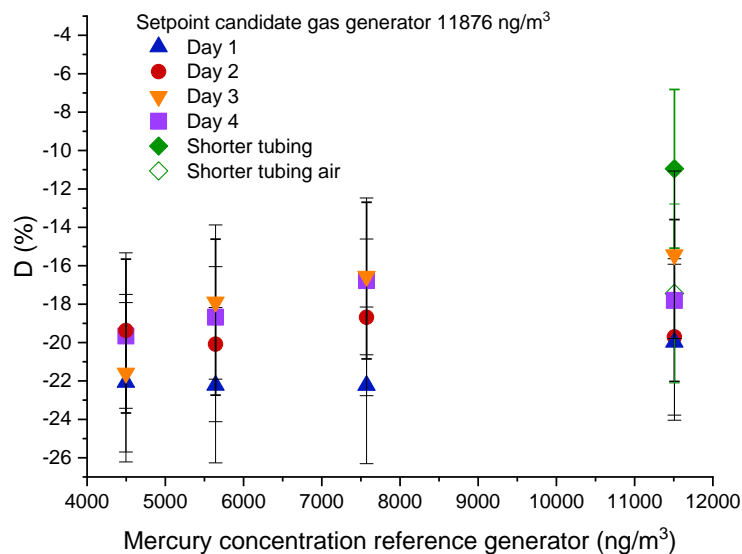


Figure 5 – Plot with the setpoints of the reference generator against the deviation of the output of the candidate generator from the setpoint compared to the reference generated. Results from the 4 measurement days and the measurements with shorter tubing with nitrogen and air as complementary gas from Channel A are shown.

The results obtained by channel A and channel B during the 4 measurement days are comparable and the average output of the candidate generator based on the results is  $(9587 \pm 393) \text{ ng m}^{-3}$  ( $k = 2$ ), with an average relative expanded uncertainty of 4.1%. The output of the generator has a deviation of  $-2289 \text{ ng m}^{-3}$  which is a relative deviation of -19 % compared to the reference generator. During the 4 measurement days the set-up was kept the same. The length of tubing from the generator to the analyser is determining the back pressure in the mercury reservoir of the candidate generator. After the 4 measurement days a shorter piece of tubing was used to connect the output of the candidate generator to the analyser. This increased the output of the generator and gave an average output of the candidate generator of  $(10525 \pm 430) \text{ ng m}^{-3}$  ( $k = 2$ ), with a relative expanded uncertainty of 4.1%. The output of the generator has a deviation of  $-1351 \text{ ng m}^{-3}$  which is a relative deviation of -11 % compared to the reference generator.

To determine the influence of the complementary gas the candidate generator was also tested with compressed air instead of nitrogen. The average output of the candidate generator operated with compressed air is  $(9753 \pm 431) \text{ ng m}^{-3}$  ( $k = 2$ ), with a relative expanded uncertainty of 4.4%. The output of the generator has a deviation of  $-2123 \text{ ng m}^{-3}$  which is a relative deviation of -18 % compared to the reference generator. This is a deviation of 7 % compared to the measurement with nitrogen as complementary gas. For the generator this should not make a difference as a critical orifice flow is independent of the complementary gas. Nevertheless, it is important to clearly state which complementary gas is used during the calibration to operate the candidate generator.

In conclusion, the generator output is stable over the measurement period of 1 month. The results of the 12 comparisons for both channel A and channel B are comparable even though the results are compared to different mercury concentrations from the reference generator. For the bracketing sequence mercury concentrations for the candidate generator and the reference standard can be selected such that the measuring system produces responses which are within  $\pm 50 \%$ . This is only possible if 1) the output of the candidate generator, 2) the output of the reference standard and 3)

the analyser responses are all linear within the selected mercury concentrations. Nevertheless, it is preferred to select mercury concentrations using the same nominal settings.

The relative deviation between the candidate generator ( $c_{cand}$ ) and the calibrated candidate generator output mercury concentration ( $c$ ) has an average value of -19 %. The output is caused by the set-up in which the output of the candidate generator is effected by the backpressure. For all 4 measurements day's the set-up was the same and the backpressure did not vary significantly. Decreasing the length of the tubing and the backpressure leads to results with an deviation of -9 %. Correction for the backpressure is not a straightforward correction as the actual flow changes across the orifice when a backpressure is applied.

Therefore, the length of tubing between the generator and the analyser has a large influence on the output of the generator. To prevent this in the future the equipment was improved to ensure the output is not influenced by backpressure created by the tubing at the outlet.

#### 4.3.2 Measurements March 2023

After the modification the generator was calibrated again at VSL. During this calibration the reference generator was used to obtain 3 mercury concentrations each concentration as compared with the output of the candidate generator, which has a fixed mercury concentration output after improvement of  $(10640 \pm 242)$  ng m<sup>-3</sup> at 1032 mbar (Table 2). This sequence was repeated 3 times with air as complementary gas and 3 times with air as complementary gas.

For calculation of the mercury concentration,  $c$ , at least triplicate responses for the mercury concentration from the candidate generator and quadruple responses for the mercury concentration from the reference standard were used. The responses were used to calculate the actual mercury concentration and the uncertainty according to the calibration protocol [2]. A zero correction was applied. No pressure correction was applied, this was no longer necessary after modification of the generator. The raw data and the results obtained can be found in a repository [13].

The RSD values of the R values are all  $\leq 2.0$  %. The mercury concentration,  $c$ , obtained with the PSA generator, the uncertainty,  $U(c)$ , of the concentration and the deviation,  $D_c$  and  $D_{rel}$ , compared to the setpoint were calculated according to the procedure in the calibration protocol using the data processing software with air as complementary gas (Table 13, Figure 6) and nitrogen as complementary gas (Table 14, Figure 6).

Table 13 – Summary of the results obtained for the mercury concentration,  $c$ , obtained with the PSA generator after modification using air as complementary gas, the uncertainty,  $U(c)$ , of the concentration and the deviation,  $D_c$  and  $D_{rel}$ , compared to the setpoint for channel A and channel B.

Measurement day	$c_{ref} \pm U$	$c_{cand}$	$c$	$U(c) (k = 2)$		Deviation	
	(ng m <sup>-3</sup> ) ( $k = 2$ )	(ng m <sup>-3</sup> )	(ng m <sup>-3</sup> )	(ng m <sup>-3</sup> )	(%)	(ng m <sup>-3</sup> )	(%)
<b>Channel A</b>							
1	13883 ± 555	10640	11414	464	4.1	774	7.3
	11477 ± 459	10640	11460	461	4.0	820	7.7
	8523 ± 341	10640	11364	474	4.2	724	6.8
2	13883 ± 555	10640	11088	446	4.0	448	4.2
	11477 ± 459	10640	11272	466	4.1	632	5.9
	8523 ± 341	10640	11228	456	4.1	588	5.5
3	13883 ± 555	10640	11490	467	4.1	850	8.0
	11477 ± 459	10640	11338	463	4.1	698	6.6

	8523 ± 341	10640	11497	469	4.1	857	8.1
<b>Measurement day</b>	<b><math>c_{ref} \pm U</math></b>	<b><math>c_{cand}</math></b>	<b><math>c</math></b>	<b><math>U(c) (k = 2)</math></b>		<b>Deviation</b>	
	( $\mu\text{g m}^{-3}$ ) ( $k = 2$ )	( $\text{ng m}^{-3}$ )	( $\text{ng m}^{-3}$ )	( $\text{ng m}^{-3}$ )	(%)	( $\text{ng m}^{-3}$ )	(%)
<b>Channel B</b>							
1	13883 ± 555	10640	11692	479	4.1	1052	9.9
	11477 ± 459	10640	11617	481	4.1	977	9.2
	8523 ± 341	10640	11441	475	4.2	801	7.5
2	13883 ± 555	10640	11190	453	4.0	550	5.2
	11477 ± 459	10640	11330	460	4.1	690	6.5
	8523 ± 341	10640	11252	455	4.0	612	5.8
3	13883 ± 555	10640	11528	466	4.0	888	8.3
	11477 ± 459	10640	11358	466	4.1	718	6.7
	8523 ± 341	10640	11518	467	4.1	878	8.3

Table 14 – Summary of the results obtained for the mercury concentration,  $c$ , obtained with the PSA generator after modification using nitrogen as complementary gas, the uncertainty,  $U(c)$ , of the concentration and the deviation,  $D_c$  and  $D_{rel}$ , compared to the setpoint for channel A and channel B.

<b>Measurement day</b>	<b><math>c_{ref} \pm U</math></b>	<b><math>c_{cand}</math></b>	<b><math>c</math></b>	<b><math>U(c) (k = 2)</math></b>		<b>Deviation</b>	
	( $\text{ng m}^{-3}$ ) ( $k = 2$ )	( $\text{ng m}^{-3}$ )	( $\text{ng m}^{-3}$ )	( $\text{ng m}^{-3}$ )	(%)	( $\text{ng m}^{-3}$ )	(%)
<b>Channel A</b>							
1	13883 ± 555	10640	11626	468	4.0	986	9.3
	11477 ± 459	10640	11712	482	4.1	1072	10.1
	8523 ± 341	10640	11766	472	4.0	1126	10.6
2	13883 ± 555	10640	11852	475	4.0	1212	11.4
	11477 ± 459	10640	11380	457	4.0	740	7.0
	8523 ± 341	10640	11320	456	4.0	680	6.4
3	13883 ± 555	10640	11431	473	4.1	791	7.4
	11477 ± 459	10640	11400	458	4.0	760	7.1
	8523 ± 341	10640	11358	461	4.1	718	6.7
<b>Measurement day</b>	<b><math>c_{ref} \pm U</math></b>	<b><math>c_{cand}</math></b>	<b><math>c</math></b>	<b><math>U(c) (k = 2)</math></b>		<b>Deviation</b>	
	( $\mu\text{g m}^{-3}$ ) ( $k = 2$ )	( $\text{ng m}^{-3}$ )	( $\text{ng m}^{-3}$ )	( $\text{ng m}^{-3}$ )	(%)	( $\text{ng m}^{-3}$ )	(%)
<b>Channel B</b>							
1	13883 ± 555	10640	11644	470	4.0	1004	9.4
	11477 ± 459	10640	11737	478	4.1	1097	10.3
	8523 ± 341	10640	11738	471	4.0	1098	10.3
2	13883 ± 555	10640	11713	470	4.0	1073	10.1
	11477 ± 459	10640	11338	457	4.0	698	6.6
	8523 ± 341	10640	11339	460	4.1	699	6.6
3	13883 ± 555	10640	11527	466	4.0	887	8.3
	11477 ± 459	10640	11424	459	4.0	784	7.4
	8523 ± 341	10640	11375	460	4.0	735	6.9

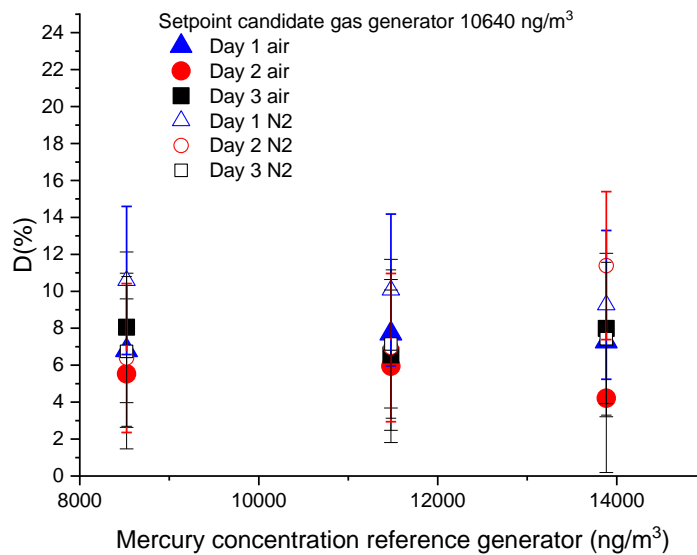


Figure 6 – Plot with the setpoints of the reference generator against the deviation of the output of the candidate generator from the setpoint compared to the reference generated. Results from the three measurement days with air as complementary gas and the three measurement days with nitrogen as complementary gas from Channel A are shown.

The candidate generator was tested with air and nitrogen as complementary gas. All measurements were repeated three times. The results obtained by channel A and channel B during the three measurements are comparable and the average output of the candidate generator with air as complementary gas is  $(11393 \pm 465) \text{ ng m}^{-3}$  ( $k = 2$ ), with an average relative expanded uncertainty of 4.08 %. The output of the generator has a deviation of  $753 \text{ ng m}^{-3}$  which is a relative deviation of 7.08 % compared to the reference generator. The results obtained with nitrogen as complementary gas give a higher output and a deviation of  $(11538 \pm 466) \text{ ng m}^{-3}$  ( $k = 2$ ), with an average relative expanded uncertainty of 4.04 %. The output of the generator has a deviation of  $898 \text{ ng m}^{-3}$  which is a relative deviation of 8 % compared to the reference generator. The results in air and nitrogen are comparable within the calculated uncertainty of 4%. As the measurements at TUV were performed with air as complementary gas the results obtained with air were used as final results.

#### 4.3.3 Measurements June 2023

After the measurements at VSL the PSA 10.532 fixed output generator was shipped to TUV. At TUV the calibrated generator was used as reference generator to calibrate another candidate generator, a PSA multipoint 10.536 elemental mercury generator. In this way a calibration chain was obtained. Both the reference generator and candidate generator were operated with air as complementary gas. The candidate generator was operated with a reservoir flow of  $1.7 \text{ mL min}^{-1}$  and a dilution flow of  $6.8 \text{ L min}^{-1}$  giving a setpoint of  $11350 \text{ ng m}^{-3}$  (Table 15). During the measurement the pressure was 1019 mbar giving a pressure corrected setpoint of  $11296 \text{ ng m}^{-3}$ .

Table 15 - Summary of the results obtained for the mercury concentration,  $c$ , obtained at TUV using air as complementary gas, the uncertainty,  $U(c)$ , of the concentration and the deviation,  $D_c$  and  $D_{rel}$ , compared to the setpoint.

$c_{ref} \pm U$	$c_{cand}$	$c'_{cand}$	$c$	$U(c) (k = 2)$		Deviation	
$(\text{ng m}^{-3}) (k = 2)$	$(\text{ng m}^{-3})$	$(\text{ng m}^{-3})$	$(\text{ng m}^{-3})$	$(\text{ng m}^{-3})$	(%)	$(\text{ng m}^{-3})$	(%)
$11393 \pm 465$	11350	11296	11422	467	4.09	126	1.1



The output of the candidate generator is  $(11393 \pm 465) \text{ ng m}^{-3}$  ( $k = 2$ ). The average relative expanded uncertainty of 4.09 % slightly increased compared to the uncertainty of the reference generator having an uncertainty of 4.08 %. The deviation compared to the corrected setpoint is  $126 \text{ ng m}^{-3}$  which is a relative deviation of 1.1%.

#### 4.3.4 Measurements September 2023

The PSA 10.532 fixed output generator was shipped back to VSL after measurements performed at TUV. At VSL the measurements from March 2023 with air as complementary gas were repeated twice (Table 16).

Table 16 – Summary of the results obtained for the mercury concentration,  $c$ , obtained with the PSA 10.532 fixed output generator after modification using air as complementary gas, the uncertainty,  $U(c)$ , of the concentration and the deviation,  $D_c$  and  $D_{rel}$ , compared to the setpoint for channel A and channel B.

Measurement day	$c_{ref} \pm U$	$c_{cand}$	$c$	$U(c) (k = 2)$		Deviation	
	$(\text{ng m}^{-3}) (k = 2)$	$(\text{ng m}^{-3})$	$(\text{ng m}^{-3})$	$(\text{ng m}^{-3})$	(%)	$(\text{ng m}^{-3})$	(%)
<b>Channel A</b>							
1	$13798 \pm 552$	10640	11072	475	4.3	432	4.1
	$11411 \pm 456$	10640	11151	472	4.2	511	4.8
	$8478 \pm 339$	10640	10974	470	4.3	334	3.1
2	$13798 \pm 552$	10640	11072	475	4.3	432	4.1
	$11411 \pm 456$	10640	11151	472	4.2	511	4.8
<b>Channel B</b>							
Measurement day	$c_{ref} \pm U$	$c_{cand}$	$c$	$U(c) (k = 2)$		Deviation	
	$(\mu\text{g m}^{-3}) (k = 2)$	$(\text{ng m}^{-3})$	$(\text{ng m}^{-3})$	$(\text{ng m}^{-3})$	(%)	$(\text{ng m}^{-3})$	(%)
1	$13798 \pm 552$	10640	11021	451	4.1	381	3.6
	$11411 \pm 456$	10640	11199	476	4.3	559	5.3
	$8478 \pm 339$	10640	11211	470	4.2	571	5.4
2	$13798 \pm 552$	10640	11021	451	4.1	381	3.6
	$11411 \pm 456$	10640	11199	476	4.3	559	5.3

The deviations are comparable within the uncertainty compared to the measurements performed in March 2023 (Figure 7).

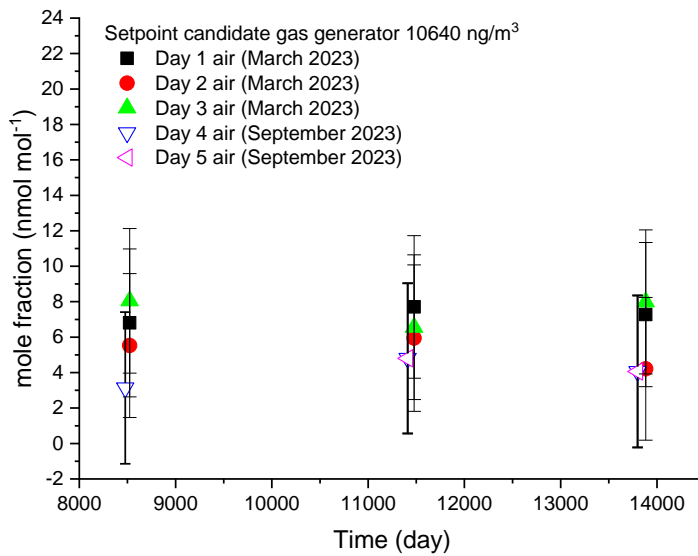


Figure 7 - Plot with the setpoints of the reference generator against the deviation of the output of the candidate generator from the setpoint compared to the reference generated. Results from the five measurement days with air as complementary gas from Channel A are shown.

The results obtained by channel A and channel B during the two measurements are comparable and the average output of the candidate generator with air as complementary gas is  $(11107 \pm 469) \text{ ng m}^{-3}$  ( $k = 2$ ), with an average relative expanded uncertainty of 4.2 %. The output of the generator has a deviation of  $467 \text{ ng m}^{-3}$  which is a relative deviation of 4 % compared to the reference generator.

## 5. Conclusion

The calibration protocol has successfully been validated and it ready for larger scale validation with gas generators available on the market this will be done in the WP3 performance evaluation of the SI-Hg project [8]. The measurements described in the protocol could be performed and the data was processed using a script to determine the output of the candidate generator and the uncertainty of the mercury concentration. Based on the validation measurements and data processing several improvements were identified for the calibration protocol (Section 6). The results of this validation and the performance evaluation were used to improve the protocol. After finalizing the protocol, it will be converted into written documentary standards in collaboration with standardisation committees such as CEN/TC 264 "Air quality" WG8 ". In three years from now this standard will be available which is essential to obtain comparable mercury measurements results to underpin global efforts to control and reduce the concentration of mercury in the environment, comply with legislation and protect human health.

## 6. Improvements for the protocol

- A note was added to section 8.4.1: NOTE: clearly state on the calibration certificate which complementary gas is used during the calibration to operate the candidate generator.
- Section 8.4.2.1: Select a mercury concentration using the same nominal setpoint for the candidate generator ( $c_{cand(i)}$ ) and the reference standard ( $c_{ref(i)}$ ) such that the measuring system produces responses which are within  $\pm 50\%$ .
- Section 8.5: NOTE: Ensure stable responses are obtained with a RSD  $\leq 2.0\%$ . The first responses for the zero measurements, reference standard and candidate generator, on both channel A and channel B if applicable, can show instability when going from one concentration level to the other.
- Section A1.2.3 For this example the reproducibility uncertainty was set to zero. Reproducibility measurements performed on 3 measurement days show that the results from Channel A and Channel B and the results from the 3 days are comparable within the uncertainty without adding a contribution for the reproducibility. Therefore, it is not necessary to add extra uncertainty due to the reproducibility of the measurements. NOTE when the measurements, especially those obtained on different days, are not comparable within the uncertainty it is advised to add uncertainty for the reproducibility to the uncertainty calculation.

## References

1. Interim EPA traceability protocol for qualification and certification of elemental mercury gas generators, July 2009, NIST [Interim EPA Traceability Protocols for the Qualification and Certification of Elemental and Oxidized Mercury Gas Generators | US EPA](#)
2. I. de Krom, A.M.H. van der Veen, F. Gugole, SI-Hg D1 Protocol for the SI-traceable calibration of elemental mercury (Hg<sup>0</sup>) gas generators used in the field, 2023, DOI: <https://zenodo.org/records/10124313>
3. H. Ent, I. van Andel, M. Heemskerk, R.P. Otterloo, W. Bavius, A. Baldan, M. Horvat, R.J.C. Brown, C.R. Quétel, A gravimetric approach to providing SI traceability for concentration measurement results of mercury vapor at ambient air levels, *Meas. Sci. Technol.*, 25 (2014), 115801, <https://doi.org/10.1088/0957-0233/25/11/115801>.
4. I. de Krom, W. Bavius, R. Ziel, E. Efremov, D. van Meer, P. van Otterloo, I. van Andel, D. van Osselen, M. Heemskerk, A.M.H. van der Veen, M.A. Dexter, W.T. Corns, H. Ent, Primary mercury gas standard for the calibration of mercury measurements, *Measurement*, 169 (2021), 108351, <https://doi.org/10.1016/j.measurement.2020.108351>.
5. I. de Krom, W. Bavius, R. Ziel, E.A. McGhee, R.J.C. Brown, I. Živkovic, J. Gacnik, V. Fajon, J. Kotnik, M. Horvat, H. Ent, Comparability of calibration strategies for measuring mercury concentrations in gas emission sources and the atmosphere, *Atmospheric Measurement Techniques*, 14 (2021), 2317, <https://doi.org/10.5194/amt-14-2317-2021>.
6. ISO 6145-8:2005: Gas analysis – Preparation of calibration gas mixtures using dynamic volumetric methods – Part 8: Diffusion method.
7. Federica Gugole, A.M.H van der Veen, SI-Hg data processing script, 2023, [vanswindenlaboratory/SI-Hg: v1.0 release \(zenodo.org\)](#) DOI: 10.5281/zenodo.8426350.
8. I. de Krom, E. Efremov, A.M.H. van der Veen, C. Röllig, S. Simon, SI-Hg D5 Report on the performance evaluation of at least three elemental mercury gas generators on the market, 2023.
9. T. D. Andron, W. T. Corns, I. Živković, S. W. Ali, S. Vijayakumaran Nair, M. Horvat, Traceable and continuous flow calibration method for gaseous elemental mercury at low ambient concentrations, *Atmospheric Measurement Techniques*, manuscript submitted for review, 2023.
10. 22.05 Validatie kwik analyser PSA Sir Galahad II, 2022, Internal VSL report
11. ISO 6145-9:2009: Gas analysis – Preparation of calibration gas mixtures using dynamic volumetric methods – Part 9: Saturation method.
12. ISO 5725-2:2019: Accuracy (trueness and precision) of measurement methods and results – Part 2: Basic method for the determination of repeatability and reproducibility of a standard measurement method.
13. I. de Krom, Data for SI-Hg D2 validation report for the calibration of elemental mercury gas generators including information on repeatability, reproducibility and uncertainty evaluation at emission and ambient levels extended to the sub ng/m<sup>3</sup> level, 2023, <https://zenodo.org/records/8296638>