

Response to Comments by Franz Slemr (Referee #1)

The reviewer's comments are shown below in bold type. My responses are in plain type.

Title: The implications of the improved processing of Tekran signal are even more important for measurements of gaseous oxidized mercury (GOM) and particulate bond mercury (PBM) than for measurements gaseous elemental mercury (GEM). A title without mentioning specifically "elemental mercury" would thus be more appropriate.

I agree that a title without mentioning specifically "elemental mercury" would be more appropriate. I changed the title accordingly, omitting "Elemental" and "Vapor".

The author claims that the Tekran measurements even with loads ≥ 10 pg are biased low. This is probably not correct because it will depend on the calibration which uses the same default integration. If the calibration is made with 10 pg than then the measurement of 10 pg, both with the same default integration, will not be biased (the low bias of the calibration equals the low bias of the measurement and they cancel out). This should be mentioned and the text modified accordingly.

The reviewer makes a very good point that I agree should be addressed in the paper. I changed the text in several places to clarify that the Tekran method "can" yield Hg measurements that are biased low at high loading (≥ 10 pg). For instance, line 13 of Page 1 now reads "the Tekran® method can produce significant low biases", not "the Tekran method produces significant low biases". Additionally, I added the following paragraph to Sect. 3.1.1.:

"It is possible that bias introduced by the Tekran® method can be made smaller by calibrating at loadings more similar to the loadings in the samples of interest. The measurements in Figs. 3a and S5a are calibrated with SPAN loadings >10 times higher (see Figs. 2 and S3, respectively). Calibrating the measurements in Fig. S5a using the external SPANS (Fig. S3; loadings >5 times higher) yields the linear regression equation (as in Fig. S5a) $y = 0.96(1)x - 0.09(5)$ pg. The slope of the latter equation is closer to unity (though not significantly) than that of the equation derived from Fig. S5a."

I added the following sentences to Sect. 4: "Additionally, it is possible that measurement bias introduced by the Tekran® method can be made smaller by calibrating at loadings more similar to the loadings in the samples of interest. My results demonstrate a minor reduction in bias when measurements made with the 2537B instrument are calibrated at loadings that are >5 times higher rather than at loadings >10 times higher."

Page 2, line 23: A sentence describing the hardware change of the 2537B instrument would be helpful.

I agree. I added the following brief description of the hardware changes to Sect. 2, including the manufacturer's part numbers for the components that were changed: "The instrument's sample cuvette and detector bandpass filter were replaced with a mirrored cuvette and an improved bandpass filter, respectively (Ambrose et al., 2013). The cuvette and bandpass filter were obtained from Tekran® (Part Numbers 40-25105-00 and 40-25200-02, respectively)."

Page 5, lines 11-14: Three consecutive sentences start with “I estimate the uncertainty. . .”

Trying to somehow combine these sentences seems to complicate them. However, I did reword the latter two sentences to avoid some redundancy.

Page 10, 1st paragraph: This paragraph concerning the GOM and PBM measurements with the Tekran speciation system is confusing and grossly underrates the problem. GOM and PBM measurements in AMNet could be mentioned as an example: With 2 h sampling at a rate of 10 l min⁻¹ median GOM concentrations of 1.2 – 2.5 pg m⁻³ (Gay et al., ACP, 13, 11339-11349, 2013) will give loads of 1.4 – 3 pg. I.e. half of these measurements are made with loads below 3 pg (not to say anything about 8 pg mentioned in this paragraph as a minimum load to avoid large bias) and are thus substantially biased low with correspondingly poor precision. PBM measurements and other networks could be mentioned as well. As most of the existing GOM and PBM data have been obtained by Tekran speciation system it can be said, without exaggeration, that most of them are biased low. The bias adds to the intensively discussed artefacts related to the GOM and PBM sampling.

I agree that signal-processing bias/uncertainties in Tekran®-based GOM and PBM measurements can be even more severe than in GEM/THg measurements under common sampling conditions and that the text should be modified to more accurately reflect the severity of the issue. I removed the paragraph in question and replaced it with the following one:

“Most measurements of atmospheric GOM and PBM made with the Tekran® 1130/1135 Hg speciation system (Landis et al., 2002) may be significantly biased low. For instance, “typical” median GOM and PBM concentrations measured at 21 Atmospheric Mercury Network (AMNet) sites in the United States and Canada during the years 2009–2011 were in the ranges 1.2–2.5 pg/m³ and 2.5–5.0 pg/m³, respectively. The corresponding Hg loadings are 1.4–3.0 pg and 3.0–6.0 pg. Median signal processing bias, estimated as $100 \times (\text{Hg}_{\text{Tekran}} - \text{Hg}_{\text{benchmark}}) \div \text{Hg}_{\text{Tekran}}$, would be within –39% and –19%, respectively, based on the fit in Fig. 3a and within –17% and –12%, respectively, based on the fit in Fig. S5a. Similarly, median concentrations of GOM and/or PBM measured at 10 sites in Canada during the years 2002–2011 were typically <5 pg/m³ (Cole et al., 2014), corresponding with sample loadings of 3–9 pg. The corresponding median signal processing bias would be within –19% and –8%, respectively, based on the fit in Fig. 3a and within –12% and –8%, respectively, based on the fit in Fig. S5a.”

Additionally, I added the following sentence to the Abstract: “Signal processing bias can also add significantly to uncertainties in Tekran®-based gaseous oxidized mercury (GOM) and particle bound mercury (PBM) measurements, which often derive from Hg sample loadings <5 pg.” I feel that a more comprehensive review of previous GOM and PBM measurements is beyond the scope of this paper.

Page 10, paragraph starting at line 14: Longer sampling times are another (hardware) way to increase the loads and thus reduce the biases and improve the precision especially of the GEM and TGM measurements. E.g. GEM measurements at stations in the southern hemisphere such as at Cape Point and Amsterdam Island are for these reasons already being made with 15 min sampling time which at a sampling rate of 1 l min⁻¹

yields loads of ~15 pg. For GOM and PBM measurements longer sampling times may not be feasible because of sampling artefacts and poor temporal resolution.

I agree and added the following sentence to Sect. 4: “Measurement bias and precision can in principle both be improved by employing longer sample pre-concentration times and/or higher sample flow rates to achieve higher sample loadings.” I feel it is beyond the scope of this paper to attempt to address all of the potential drawbacks to operating the Tekran® instrument with longer-than-normal pre-concentration times and/or higher-than-normal sample flow rates.

Fig. 2: There is hardly any contrast between the colours for methods VIm,a and VIa,a.

I changed the color of the VI_{a,a} series to improve the contrast against the VI_{m,a} series in Figs. 2 and S3.

10 Corrections:

I found that the fit in Fig. 3a in the main manuscript included the data points for thermal desorption peaks not detected by the Tekran® method. Those points should be excluded. The fit parameters do differ depending on whether those points are included in the fit. For instance, when those points are included, the equation of the fit in Fig. 3a is $y = 0.963(7) - 0.24(3) \text{ pg}$ ($n = 152$, $r^2 = 0.998$). When those points are excluded, the fit equation is $y = 1.00(1) - 0.45(8) \text{ pg}$ ($n = 74$, $r^2 = 0.997$). This change does not substantially influence my conclusions.

In the revised manuscript, the fit in Fig. 3a excludes data points for TD peaks undetected by the Tekran® method. Accordingly, I revised Fig. 3b, Fig. 3c, Tekran® bias values in Table 1, Tekran® signal processing uncertainty values in Table 2, and other affected numerical values in the text.

20

25

30

Response to Comments by Anonymous Referee #3

Although the reviewer recommended that the manuscript be published “as is”, I am providing responses the reviewer’s comments here. The reviewer’s comments are shown below in bold type. My responses are in plain type.

- 5 **Supports previous work identifying the need for external peak integration at low load levels. Manually identification of start and/or end peaks is impractical and could be biased. Use automated (Vaa) identification**

By “impractical”, I assume the reviewer is referring to the extra time needed to carry out manual peak identification, as opposed to automated peak identification. Automatic peak identification is certainly the most
10 time efficient, but the performance of any automated method must be evaluated as I have done. Regarding bias, I would argue that it can be minimized with training. It is important to treat all peaks consistently so that any relative bias cancels out during calibration.

Uncertain why Tekran LOD’s were calculated using 2.5 liter sample volumes when most sample 5 liter volumes.

- 15 The mass-based LODs I calculated are independent of sample volume. At a given Hg concentration, higher sample volumes will be required to achieve higher Hg loadings.

20

25

30

Improved Methods for Signal Processing in Measurements of ~~Elemental Mercury Vapor~~ by Tekran® 2537A and 2537B Instruments

Jesse L. Ambrose¹

¹College of Engineering and Physical Sciences, University of New Hampshire, Durham, 03824, USA

5 Correspondence to: Jesse L. Ambrose (Jesse.Ambrose@unh.edu)

Abstract. Atmospheric Hg measurements are commonly carried out using Tekran® Instruments Corporation's model 2537 Hg vapor analyzers, which employ gold amalgamation pre-concentration sampling and detection by thermal desorption/atomic fluorescence spectrometry. A generally overlooked and poorly characterized source of analytical uncertainty in those measurements is the method by which the raw Hg atomic fluorescence signal is processed. Here I describe new software-based methods for processing the raw signal from the Tekran® 2537 instruments, and I evaluate the performances of those methods together with the standard Tekran® internal signal processing method. For test datasets from two Tekran® instruments (one 2537A and one 2537B), I estimate that signal processing uncertainties in Hg loadings determined with the Tekran® method are within $\pm[1\% + 1.2 \text{ pg}]$ and $\pm[6\% + 0.21\text{--}0.94 \text{ pg}]$, respectively. I demonstrate that the Tekran® method can produces significant low biases ($\geq 5\%$) not only at low Hg sample loadings ($< 5 \text{ pg}$) but also at tropospheric background concentrations of gaseous elemental mercury (GEM) and total mercury (THg) (~ 1 to 2 ng/m^3) under typical operating conditions (sample loadings of $5\text{--}10 \text{ pg}$). Signal processing uncertainties associated with the Tekran® method can therefore represent a significant unaccounted for addition to the overall ~ 10 to 15% uncertainty previously estimated for Tekran®-based GEM and THg measurements. Signal processing bias can also add significantly to uncertainties in Tekran®-based gaseous oxidized mercury (GOM) and particle bound mercury (PBM) measurements, which often derive from Hg sample loadings $< 5 \text{ pg}$. In comparison, estimated signal processing uncertainties associated with the new methods described herein are low, ranging from within $\pm 0.053 \text{ pg}$, when the Hg thermal desorption peaks are defined manually, to within $\pm[2\% + 0.080 \text{ pg}]$ when peak definition is automated. Mercury limits of detection decreased by ~ 31 to 88% when the new methods are used in place of the Tekran® method. I recommend that signal processing uncertainties be quantified in future applications of the Tekran® 2537 instruments.

Formatted: Superscript

1 Introduction

25 Gold amalgamation pre-concentration, followed by thermal desorption (TD) in Ar carrier gas and detection via atomic fluorescence spectrometry (AFS) is a commonly used method for quantifying atmospheric elemental mercury vapor, $\text{Hg}^0(\text{g})$ (hereafter referred to as gaseous elemental mercury, GEM) (Schroeder et al., 1995; Gustin and Jaffe, 2010; Pandey et al., 2011). Coupled with various sample capture and pre-treatment methods, the above measurement scheme is also used for quantitative analysis of atmospheric gaseous oxidized mercury (GOM) (Stratton and Lindberg, 1995; Landis et al., 2002; Lyman et al., 2007), total gaseous mercury (TGM \equiv GEM + GOM) (Ambrose et al., 2015), atmospheric particle-bound mercury (PBM) (Landis et al., 2002), atmospheric total mercury (THg \equiv GEM + GOM + PBM) (Jaffe et al., 2005), and total aqueous Hg

(USEPA, 2002). Most AFS-based atmospheric Hg measurements employ Tekran® Instruments Corporation’s model 2537 Hg vapor analyzers (versions A and B; hereafter referred to collectively as ‘the Tekran® analyzer’) (Schroeder et al., 1995; Landis et al., 2002; Tekran Corporation, 2006, 2007).

Several previous reports identified analytical uncertainties as critical in limiting scientific understanding of environmental Hg cycling (Jaffe and Gustin, 2010; Pirrone et al., 2013; Jaffe et al., 2014). An often overlooked source of analytical uncertainty is the method by which the raw Hg atomic fluorescence (AF) signal is processed by the Tekran® analyzer. Although most researchers rely on the Tekran® analyzer’s embedded software to automatically integrate the Hg TD peaks, Swartzendruber et al. (2009a) and Slemr et al. (2016) demonstrated that the accuracy and precision of the Tekran® peak integration method can significantly decline at low Hg loadings.

To better characterize analytical uncertainty associated with atmospheric Hg measurements made with the Tekran® analyzer, and in an attempt to improve upon existing measurement methods for atmospheric Hg, I developed new software-based methods for offline processing of the raw Hg AF signal from the Tekran® analyzer. Here I describe the key features of the new signal processing methods and characterize their performances, together with that of the standard Tekran® signal processing method.

2 Experimental

Using National Instruments™ LabVIEW software (version 12.0), I developed a virtual instrument (VI) for offline processing of the serial data output (“RAWDUMP” format) from Tekran® models 2537A and 2537B mercury vapor analyzers. The VI is characterized and validated using data collected as part of the Nitrogen, Oxidants, Mercury, and Aerosol Distributions, Sources, and Sinks (NOMADSS) campaign (https://www.eol.ucar.edu/field_projects/nomadss) using a Tekran® 2537A instrument and a 2537B instrument incorporated in the University of Washington’s Detector for Oxidized Hg Species (Swartzendruber et al., 2009b; Lyman and Jaffe, 2012; Ambrose et al., 2013, 2015). The instrument configurations are described in Ambrose et al. (2015). Instrument operating parameters are given in Table S1 in the Supporting Information. The most significant difference between the two instruments tested is that hardware in the 2537B instrument was modified to improve its signal-to-noise ratio. The instrument’s sample cuvette and detector bandpass filter were replaced with a mirrored cuvette and an improved bandpass filter, respectively (Ambrose et al., 2013). The cuvette and bandpass filter were obtained from Tekran® Instruments (Part Numbers 40-25105-00 and 40-25200-02, respectively).

All linear regression equations reported herein were calculated by the Bisquare method (unweighted) in LabVIEW. All stated uncertainties represent 95% confidence intervals, unless otherwise specified.

2.1 Virtual Instrument Design Overview

2.1.1 Operation

The VI's Hg atomic fluorescence signal processing method parses the Tekran[®] analyzer's serial data output ("RAWDUMP" format; see Sect. S1 in the Supporting Information) using text delimiters associated with each component (Figs. S1 and S2 in the Supporting Information).

At the start of an analysis, the VI computes the overall mean baseline standard deviation, σ_{bl} , for all samples in the data file to be analyzed. The baseline standard deviation for each sample is first calculated as the 10 point (1 s) running mean of the baseline measurements made from 1 s after the start of the AF signal recording interval to 2 s prior to the approximate start of the Au trap desorption cycle (*i.e.*, the "BL time" specified by the operator in step 2 below). The value of σ_{bl} is then calculated from the values for all individual samples.

For each sample analysis cycle, the VI carries out the following operations:

1. The "Raw data" string (Fig. S1 in the Supporting Information) is converted to an array of 10 Hz AF signal values.
2. The timestamp, cycle type flag, and trap identity are extracted from the "Final data" string (Fig. S1 in the Supporting Information).
3. An x - y plot is generated from the data array created in step 1.
4. Unless the VI is set to automatically define the Hg thermal desorption peaks, the user is prompted to identify the start time, t_{start} , of the TD peak. This selection is accomplished by first manually setting the placement of a cursor on the x - y plot generated in step 3 and then using a control to extract the associated coordinates. Alternatively, values of t_{start} can be set to values defined automatically during the "initialization" procedure (see below). The mean baseline voltage at the start of the TD peak is calculated over the interval from t_{start} to $t_{start} - 9$ ds ($n = 10$ data points).
5. The coordinates of the Hg TD peak maximum are identified automatically from the data array created in step 1. In my experience, the Tekran[®] analyzer's baseline generally decreases (slopes downward) across the Hg TD peak. I parameterized the VI for such a condition by identifying the TD peak maximum as the largest AF signal value recorded after t_{start} .
6. A preliminary TD peak height value is calculated from the mean baseline voltage at the start of the peak (step 4) and the maximum voltage (step 5). In cases when the preliminary peak height is a negative number, the VI sets the value to σ_{bl} . (See Sect. S3 in the Supporting Information for further details.)
7. Based on user-defined settings on the VI's front panel, the TD peak end coordinates are either selected manually (as for the peak start time in step 4) or calculated automatically from the peak maximum in step 5 and the initialization parameters (as described below).

8. The baseline beneath the Hg TD peak is calculated. For this purpose, the baseline at the beginning of the peak is defined by the coordinates of t_{start} (step 4) and the preceding nine AF signal values. The baseline at peak end is similarly defined by the peak end coordinates in step 7 and the trailing nine AF signal values. The baseline coordinates beneath the peak are calculated by linear regression ($n = 20$ data points).
9. The baseline standard deviation at the Hg TD peak is estimated from the mean residual of the regression in step 8 (see Sect. S4 in the Supporting Information for further details). I denote this value as “ $\sigma_{\text{bl, fit}}$ ” to differentiate it from σ_{bl} defined above.
10. The baseline-slope-corrected Hg TD peak height is calculated from the peak maximum voltage (step 5) and the calculated baseline voltage at the time of the peak maximum (step 8).

At the end of an analysis, an output data file is created, which contains, among other parameters, the baseline-slope-corrected Hg TD peak height (step 10) for each sample in the data file that was processed.

Hereafter I refer to three different configurations of the VI: “VI_{m,m}” when the peak start and end times are both identified manually; “VI_{m,a}” when only the peak start time is identified manually, and “VI_{a,a}” when the peak start and end times are both identified automatically. I abbreviate the Tekran[®] peak integration method as “the Tekran[®] method”.

2.1.2 Initialization

The VI_{m,a} and VI_{a,a} methods are “initialized” by fitting eq. 1 to the Hg thermal desorption peaks recorded for a pair of calibration gas analysis cycles (*i.e.*, one A SPAN and one B SPAN; example fits are shown in Fig. 1 and in Fig. S43 in the Supporting Information).

$$S(t) = (A \times e^{b \times t}) + S_{\text{offset}} \quad (1)$$

Here, $S(t)$ is the time-dependent 10 Hz Hg fluorescence signal along the tail of the Hg TD peak, A is the peak amplitude (which is approximately equal to the peak height, H), b is the decay constant, S_{offset} is the baseline offset (*i.e.*, the difference between $\lim[S(t)]_{t \rightarrow \infty}$ and zero), and t is expressed in units of deciseconds (ds). The fit includes 150 $S(t)$ values, starting with the peak maximum, S_{max} (Fig. 1 and Fig. S4 in the Supporting Information). The values of S_{offset} and A are constrained to the baseline minimum, S_{min} , and $S_{\text{max}} - S_{\text{offset}}$, respectively. (For the data shown in Fig. 1, the values of S_{offset} and A are approximately equal to 70 and 170 mV, respectively.)

The two decay constants (one for each Au trap) are the key parameters derived from the initialization procedure (see Sect. 2.1.3). I estimate the uncertainty in each b value from the linear sum of two terms: (1) the relative difference between unity and the slope of a linear regression fit to a plot of the calculated versus measured Hg TD peak decays (Fig. 1b and Fig. S4b in the Supporting Information); and (2) the 95% confidence interval (~~CI~~) in the slope of the fit. For the data shown in Fig. 1, the uncertainty in b is estimated to be 9%.

During initialization, peak start times, t_{start} , are also calculated for the pair of calibration gas analysis cycles. For this purpose, the VI defines t_{start} as the first 10 Hz Hg fluorescence signal value in a series of seven consecutively increasing signal values.

If the operator chooses (via a control on the VI's front panel) to automatically define the Hg TD peak start times, the values of t_{start} calculated during initialization are assigned (paired by Au trap identity) to all samples in the data file to be processed.

2.1.3 Automatic Determination of the Hg Thermal Desorption Peak End Time

It is necessary to define each Hg thermal desorption peak's end time, t_{end} , within the interval during which the Tekran[®] analyzer's Hg atomic fluorescence signal is recorded (38.9 s in Fig. 1a and Fig. S4a in the Supporting Information). Therefore, for each sample the VI defines t_{end} as the time at which $S(t)$ decays to a value equal to, or less than, a fraction, f , of the amplitude, A_{SPAN} (derived as in Sect. 2.1.2), determined from a calibration gas analysis cycle on the same Au trap:

$$S(t_{\text{end}}) \leq (f \times A_{\text{SPAN}}) + S_{\text{offset}} \quad (2)$$

Substituting the right-hand side of eq. 2 for $S(t)$ in eq. 1, and solving for $t (= t_{\text{end}})$ yields an analytical expression for the upper-bound value of t_{end} :

$$t_{\text{end}} = \ln \left(f \times \frac{A_{\text{SPAN}}}{A_i} \right) \times b^{-1} \cong \ln \left(f \times \frac{A_{\text{SPAN}}}{H_i} \right) \times b^{-1} \quad (3)$$

Here, A_i and H_i are the TD peak amplitude and the initial peak height (from step 6 in Sect. 2.1.1), respectively, of the sample for which the baseline-slope-corrected peak height is to be quantified. The partial equality in eq. 3 reflects the fact that $A_{\text{SPAN}}/A_i \cong H_{\text{SPAN}}/H_i$ and $A_{\text{SPAN}} \cong H_{\text{SPAN}}$. The Au-trap-dependent value of the decay constant, b , is derived as in Sect. 2.1.2.

The value of f is chosen such that t_{end} is ≥ 0 at the smallest expected value of H_i , which the VI estimates from σ_{bl} (Sect. 2.1.1). The upper-bound value of f is estimated separately for each Au trap by solving eq. 3 for f , with $t_{\text{end}} = 0$ and $H_i = \sigma_{\text{bl}}$:

$$f = \left(\frac{\sigma_{\text{bl}}}{A_{\text{SPAN}}} \right) \quad (4)$$

The Au trap-dependent f values are then averaged prior to application in eq. 3. I estimate the uncertainty in f to be twice the standard deviation in the mean for the two Au traps. ~~UI-estimate uncertainty~~ in the value of H_i is estimated to be equal to twice the baseline standard deviation calculated in step 9 of Sect. 2.1.1. ~~UI-estimate uncertainty~~ in the automatically derived value of t_{end} , δt_{end} , is estimated by propagating ~~estimated~~ uncertainties in b , f , and H_i through eq. 3 (see below).

It is necessary to further constrain t_{end} such that for large TD peaks (*e.g.*, those recorded for SPAN samples) the automatically determined value of t_{end} is at least 10 ds before the upper bound time, t_n , of the interval during which the instrument's Hg AF signal is recorded. The VI therefore defines t_{end} as the smaller of the result of eq. 3 and t_{n-10} , where n represents the number of AF signal values recorded ($n = 389$ in Fig. 1a and Fig. S4a in the Supporting Information). A minimum value of 10 ds is also prescribed for t_{end} such that for very small TD peaks the automatically determined peak end does not occur at or before the peak max time (*e.g.*, when $f \geq H/A_{\text{SPAN}}$ in eq. 3). (See Sect. S3 in the Supporting Information for further details.)

2.1.4 Evaluation

I evaluate the performances of the VI_{m,a} and VI_{a,a} methods by applying both methods to laboratory data collected with two Tekran[®] analyzers (one model 2537A and one model 2537B). One dataset is processed for each analyzer. The test dataset collected with the 2537A instrument is shown in Fig. 2. (The test dataset for the 2537B instrument is shown in Fig. S3.)

I consider manual definition of the Hg thermal desorption peaks (the VI_{m,m} method) to be the benchmark for signal processing accuracy, and I assess the accuracies of the VI_{m,a} and VI_{a,a} methods by comparing Hg sample loadings derived from both methods with loadings derived from the VI_{m,m} method. A similar comparison is used to evaluate the accuracy of the Tekran® method. The performances of all methods are further evaluated and compared based on the Hg limits of detection (LODs) they achieve.

3 Results and Discussion

3.1 Performance Evaluations

3.1.1 The Tekran® Hg Thermal Desorption Peak Integration Method

For Hg loadings derived from the Tekran® method, Hg_{Tekran}, I define absolute bias as Hg_{Tekran} – Hg_{benchmark}, where “Hg_{benchmark}” represents loadings derived from the VI_{m,m} method. I define relative bias as $100 \times (\text{Hg}_{\text{Tekran}} - \text{Hg}_{\text{benchmark}}) \div \text{Hg}_{\text{benchmark}}$. As illustrated in Fig. 3 (and Fig. S5 in the Supporting Information), Hg loadings derived using the Tekran® method tend to be biased low, with the relative bias becoming more negative with decreasing loading. I present in Table 1 relative bias values from Fig. 3b at several discrete Hg sample loadings (analogous results for the 2537B dataset are presented in Table S2 in the Supporting Information). The corresponding Hg concentrations (ng/m³) under typical operating conditions for the Tekran® analyser (5 liter sample volumes) are also shown. My results are consistent with those of Swartzendruber et al. (2009) and Slemr et al. (2016), but also demonstrate that the Tekran® method can produce significant low biases (≥5%; see Tables 1 and S2) at tropospheric background GEM and THg concentrations (~1 to 2 ng/m³; sample loadings of 5–10 pg under typical Tekran® operating conditions).

To further characterize the performance of the Tekran® method, Fig. 4 compares Hg atomic fluorescence baselines calculated by the Tekran® method and by my VI-based peak height determination methods, for samples with Hg loadings of approximately 0.5, 1, 2, and 4 times the estimated ~0.8 pg LOD achieved with the Tekran® method (see below). (See Sect. S6 of the Supporting Information for details on how I reproduced the baselines calculated by the Tekran® method.) Figure 4 illustrates the tendency of the Tekran® method to truncate the Hg thermal desorption peaks. The peaks tend to become more severely truncated as they become smaller, and as a result, the relative biases in the corresponding Hg loadings tend to become more negative as the loadings decrease, as shown in Fig. 3b and Fig. S5b in the Supporting Information (see also Fig. 2 in Slemr et al., 2016).

It is possible that bias introduced by the Tekran® method can be made smaller by calibrating at loadings more similar to the loadings in the samples of interest. The measurements in Figs. 3a and S5a are calibrated with SPAN loadings >10 times higher (see Figs. 2 and S3, respectively). Calibrating the measurements in Fig. S5a using the external SPANS (Fig. S3; loadings >5 times higher) yields the linear regression equation (as in Fig. S5a) $y = 0.96(1)x - 0.09(5)$ pg. The slope of the latter equation is closer to unity (though not significantly) than that of the equation derived from Fig. S5a.

Formatted: Font: Italic

Formatted: Font: Italic

Table 2 presents Hg LODs for the Hg fluorescence signal processing methods and datasets I tested. The nominal Hg limit of detection of the Tekran[®] analyzer is 0.5 pg (see Sect. S7 of the Supporting Information for further details). However, some Hg thermal desorption peaks in the 2537A dataset are undetected by the Tekran[®] method for Hg loadings ≤ 0.8 pg (Figs. 2 and 3a). My results suggest that the actual Hg LOD achieved with the Tekran[®] method is ~60% higher than the nominal value.

By comparison, the LOD achieved with the VI_{m,m} method (estimated as twice the standard deviation of blank samples, $n = 62$) is 0.10 pg.

For the 2537B dataset, all Hg thermal desorption peaks are detected by the Tekran[®] method for Hg loadings > 0.2 pg (Figs. S3 and S5 in the Supporting Information), suggesting the LOD is ~60% lower than the nominal value. The lower Hg LOD achieved with the Tekran[®] method when applied to the 2537B dataset is attributable to the hardware modifications that were made to the 2537B instrument to increase its signal-to-noise ratio (Sect. 2 Ambrose et al., 2013). The 0.085 pg LOD achieved with the VI_{m,m} method is 56% lower than the Tekran[®]-based LOD (Table 2).

3.1.2 The Automated VI-Based Hg Thermal Desorption Peak Height Determination Method

By comparison with the Tekran[®] method, the VI_{a,a} method usually identifies the Hg thermal desorption peak baseline with good accuracy, even at Hg loadings near to or below the LOD achieved with the Tekran[®] method (Fig. 4). As a result, most Hg loadings derived from the VI_{a,a} method are quite accurate (Figs. 2 and 5; Table 1). For samples with Hg loadings below the estimated 0.8 pg LOD of the Tekran[®] method, the mean absolute and relative unsigned biases in the loadings derived from the VI_{a,a} method are 0.028 ± 0.005 pg and $15 \pm 3\%$, respectively ($n = 78$). The biases are very small and much smaller in magnitude than those for the Tekran[®] method (estimated at -0.80 pg and -100%). Based on the equations of the linear regressions shown in Figs. 3a and 5a (Hg loadings ≤ 10 pg), the VI_{a,a} method achieves $\geq 95.94\%$ reduction in absolute unsigned bias in calculated Hg loadings when compared with the Tekran[®] method.

The Hg LOD achieved with the VI_{a,a} method is 0.12 pg (estimated as twice the standard deviation of blank values; $n = 62$). That is value is 20% higher than the LOD achieved with the VI_{m,m} method, but 85% lower than the LOD achieved with the Tekran[®] method (Table 2). Similarly, the width of the residual distribution in Fig. 5c is 78.84% narrower than that in Fig. 3c, which reflects the improved analytical precision achieved with the VI_{a,a} method, in comparison with the Tekran[®] method.

Evaluation of the VI_{a,a} method as in Fig. 5 for the 2537B dataset (Fig. S65, Table S2 in the Supporting Information) yields the linear regression equation $y = 1.015(2)x - 0.017(8)$ pg ($r^2 = 0.9998$, $n = 132$). For samples with Hg loadings below 0.8 pg, the observed mean absolute and relative unsigned biases in the loadings derived from the VI_{a,a} method are 0.031 ± 0.007 pg and $33 \pm 9\%$, respectively ($n = 41$). The VI_{a,a} method yields larger relative bias when applied to the 2537B dataset than when applied to the 2537A dataset. However, absolute unsigned biases are equivalent (at the 95% confidence interval) for the two datasets.

Based on the equations of the linear regressions shown in Figs. S5a and S6a in the Supporting Information, the VI_{a,a} method achieves $\geq 82\%$ reduction in absolute unsigned bias in calculated Hg loadings when compared with the Tekran[®] method. The Hg LOD is 0.14 pg (Table 2), which is 59% higher than the value achieved with the VI_{m,m} method, but 31% lower than the value achieved with the Tekran[®] method. The VI_{a,a} method appears to yield a comparable improvement in analytical precision

over the Tekran[®] method when applied to the 2537A and 2537B datasets (the width of the residual distribution in Fig. S6c is 72% narrower than that in Fig. S5c).

3.1.3 The Semi-Automated VI-Based Peak Height Determination Method

The VI_{a,a} method poorly identifies the start of the Hg thermal desorption peak for some blank samples with the lowest Hg loadings. As a result, the Hg TD peak height and the calculated Hg loading tend to be underestimated for those samples (Fig. 2; Fig. S3 in the Supporting Information).

I developed the VI_{m,a} method as a compromise between the Hg TD peak definition accuracy achieved with the VI_{m,m} method and the data processing speed achieved with the VI_{a,a} method. (Peak height determination for a single sample requires from <1 second with the VI_{a,a} method to several seconds for the VI_{m,m} method; the data processing time for the VI_{m,a} method is approximately half that for the VI_{m,m} method.) Comparison of the VI_{m,a} and VI_{a,a} results in Fig. 2 (and Fig. S3 in the Supporting Information) shows that defining t_{start} manually instead of automatically yields more accurate measurements for samples with the lowest Hg loadings.

Evaluation of the VI_{m,a} method as in Fig. 5 yields the linear regression equation $y = 1.001(1)x + 0.004(3)$ pg ($r^2 = 0.99998$, $n = 152$). Comparison between the latter equation and that of the regression in Fig. 5a suggests that biases in Hg loadings derived from the VI_{a,a} and VI_{m,a} methods are not significantly different over the range of loadings in the 2537A dataset (see also Table 1). However, biases in Hg loadings derived from the VI_{m,a} method are lower at low loadings than those determined for the VI_{a,a} method (Sect. 3.1.2): for samples with Hg loadings below 0.8 pg, the mean absolute and relative unsigned biases in the loadings derived from the VI_{m,a} method are 0.012 ± 0.003 pg and $6 \pm 1\%$, respectively ($n = 78$). The estimated 0.11 pg Hg LOD achieved with the VI_{m,a} method falls in between the values achieved with the VI_{m,m} and VI_{a,a} methods (Table 2).

The VI_{m,a} method performs consistently better for the 2537B dataset than does the VI_{a,a} method. The equation of the linear regression (as in Fig. 5a) is $y = 1.005(1)x + 0.009(5)$ pg ($r^2 = 0.99994$, $n = 132$). For samples with Hg loadings below 0.8 pg, the observed mean absolute and relative unsigned biases in the loadings derived from the VI_{m,a} method are 0.010 ± 0.003 pg and $6 \pm 2\%$, respectively ($n = 41$). The estimated LOD achieved with the VI_{m,a} method is 0.10 pg and falls between the LODs achieved with the VI_{m,m} and VI_{a,a} methods (Table 2).

3.1.4. Sensitivity Analyses and Uncertainties

I estimate signal-processing uncertainties in the Hg thermal desorption peak heights derived from the VI_{m,m} method to be equal to twice the mean baseline standard deviation (σ_{bl} ; Sect. 2.1.1), corresponding with a Hg loading of 0.053 pg for the dataset in Fig. 2 (0.030 pg for the dataset in Fig. S3 in the Supporting Information). I estimate signal-processing uncertainties in the Hg loadings derived from the Tekran[®], VI_{m,m} and VI_{m,a} methods as the sum of the biases in those methods (derived as in Fig. 3) and the resulting increase (relative to the VI_{m,m} method) in the Hg limit of detection (Table 2). For the VI_{a,a} and VI_{m,a} methods, I estimate a conservative threshold uncertainty value of $\pm 2\sigma_{\text{bl}}$.

To test the sensitivity to initialization parameters of Hg loadings derived from the $VI_{a,a}$ and $VI_{m,a}$ methods I recalculated those loadings after making a series of modifications to the method initialization parameters. After each modification I recalculated the parameters of the Hg regression (*e.g.*, Fig. 5a). The modifications tested for the 2537A dataset (described further in Tables S3 and S4 in the Supporting Information) include shifts in t_{end} by $\pm\delta t_{end}$ (as defined in Sect. 2.1.3), shifts in t_{start} to the values determined (as described in Sect. 2.1.2) for the second pair of SPAN samples in Fig. 2 (applicable only to the $VI_{a,a}$ method), and initialization of the VI using the second pair of SPAN samples. (Details on the sensitivity tests applied to the 2537B dataset are described in Tables S5 and S6 in the Supporting Information.)

The above modifications result in insignificant changes (at the 95% confidence interval) to the bias parameters derived (as in Fig. 5a) for the $VI_{a,a}$ method and the 2537 dataset (Table S3 in the Supporting Information). Two sensitivity tests applied to the $VI_{a,a}$ method and the 2537B dataset result in significant changes (at the 95% confidence interval) to the calculated bias parameters (Table S5 in the Supporting Information). Results obtained with the $VI_{a,a}$ method therefore appear to be sensitive to initialization parameters in some cases, although signal processing uncertainties remain quite low and well below those estimated for the Tekran[®] method (Table 2). It is possible that a large shift in t_{start} for SPAN samples over the course of an analysis would increase the sensitivity of the $VI_{a,a}$ method to initialization parameters. Results obtained with the $VI_{m,a}$ method are insensitive (at the 95% confidence interval) to initialization parameters (Tables S4 and S6 in the Supporting Information).

I estimate “signal-processing” uncertainties in the Hg thermal desorption peak heights derived from the $VI_{m,m}$ method to be equal to twice the mean baseline standard deviation (σ_{bl} , Sect. 2.1.1), corresponding with a Hg loading of 0.053 pg for the dataset in Fig. 2 (0.030 pg for the dataset in Fig. S3 in the Supporting Information). I estimate signal processing uncertainties in the Hg loadings derived from the Tekran[®], $VI_{a,a}$, and $VI_{m,a}$ methods as the sum of the biases in those methods (derived as in Fig. 3) and the resulting increase (relative to the $VI_{m,m}$ method) in the Hg limit of detection (Table 2). For the $VI_{a,a}$ and $VI_{m,a}$ methods, I estimate a conservative threshold uncertainty value of $\pm 2\sigma_{bl}$.

~~I estimate that the overall~~ “signal processing” uncertainty attributable to my $VI_{a,a}$ method, as applied to the 2537A dataset, is ~~estimated to be~~ within $\pm[0.2\% + 0.053 \text{ pg}]$. (The first term in the latter expression represents the slope of the regression in Fig. 5a; the second term represents $2\sigma_{bl}$, which in this case is larger than the sum of the $0.000 \pm 0.006 \text{ pg}$ intercept of the fit in Fig. 5a and the 0.02 pg difference in Hg LODs achieved with the $VI_{a,a}$ and $VI_{m,m}$ methods). ~~Estimated~~ ~~I estimate that overall~~ signal processing uncertainty attributable to the $VI_{m,a}$ method is also within $\pm[0.2\% + 0.053 \text{ pg}]$. By comparison, ~~overall~~ signal processing uncertainty attributable to the Tekran[®] method (~~considering bias quantified in Sect. 3.1.1 and the estimated Hg LOD~~) is within $\pm[14\% + 10.294 \text{ pg}]$. The above uncertainty results, together with estimated Hg LODs and analogous results for the 2537B dataset, are summarized in Table 2. For both test datasets, signal processing uncertainties estimated for my VI-based methods are significantly lower than those for the Tekran[®] method. Signal processing uncertainty ranks as follows for the VI-based methods: $VI_{m,m} \leq VI_{m,a} < VI_{a,a}$.

4. Conclusions and Implications

I describe three improved methods for processing the raw Hg atomic fluorescence signal from Tekran[®] 2537A and 2537B Hg vapor analyzers. The methods incorporate manual, semi-automated, or fully automated Hg thermal desorption peak identification processes. I implement my methods through a Virtual Instrument in National Instruments[™] LabVIEW and evaluate them, together with the Tekran[®] internal Hg TD peak integration method, using test datasets from two Tekran[®] instruments (one 2537A and one 2537B).

Consistent with previous work (Swartzendruber et al., 2009; Slemr et al., 2016), my results demonstrate that Hg loadings derived from the Tekran[®] method tend to be biased low, with the relative bias becoming more negative with decreasing loading. It follows that the magnitude of the bias in Tekran[®]-based Hg measurements will depend significantly on sampling conditions (e.g., sample concentration and volume). Therefore, I recommend that signal processing bias be examined, and associated uncertainties be quantified, in all future applications of the Tekran[®] instruments, regardless of sampling arrangement.

With respect to atmospheric GEM- and THg measurements, my results demonstrate that the Tekran[®] method can produce significant low biases ($\geq 5\%$) at background concentrations (~ 1 to 2 ng/m^3) under typical operating conditions (Hg loadings of $5\text{--}10 \text{ pg}$). My results therefore indicate that post-processing the raw Tekran[®] data ~~can will~~ yield significant improvements in the accuracy of the derived Hg concentrations under much broader environmental conditions than previously recognized. Such conditions should not be assumed to be limited to those where GEM concentrations can be significantly depleted, as can occur in the free troposphere and lower stratosphere (e.g., Talbot et al., 2007; Lyman and Jaffe, 2012; Timonen et al., 2012; Gratz et al., 2015), and at the surface in polar and mid-latitude regions under special photochemical conditions (e.g., Schroeder et al., 1998; Obrist et al., 2011).

~~Most measurements of atmospheric GOM and PBM Under typical sampling conditions employed made~~ with the Tekran[®] 1130/1135 Hg speciation system (Landis et al., 2002); ~~with sample collection volumes on the order of 600–1,800 standard liters, Hg loadings during GOM and PBM analyses will be low ($< 5 \text{ pg}$) for ambient concentrations $< 8 \text{ pg/m}^3$ may be significantly biased low. For instance, “typical” median GOM and PBM concentrations measured at 21 Atmospheric Mercury Network (AMNet) sites in the United States and Canada during the years 2009–2011 were in the ranges $1.2\text{--}2.5 \text{ pg/m}^3$ and $2.5\text{--}5.0 \text{ pg/m}^3$, respectively. The corresponding Hg loadings are $1.4\text{--}3.0 \text{ pg}$ and $3.0\text{--}6.0 \text{ pg}$. Median signal processing bias, estimated as $100 \times (\text{Hg}_{\text{Tekran}} - \text{Hg}_{\text{benchmark}}) \div \text{Hg}_{\text{Tekran}}$, would be within -39% and -19% , respectively, based on the fit in Fig. 3a and within -17% and -12% , respectively, based on the fit in Fig. S5a. Relative biases much greater than 5% in Tekran[®]-based GOM and PBM measurements should therefore be limited to the lowest measured ambient concentrations. The distribution of GOM or PBM reduction over multiple denuder or particulate filter desorption cycles, respectively, would tend to reduce effective sample loadings and therefore exacerbate signal processing bias in applications of the Tekran[®] speciation system. Similarly, median concentrations of GOM and/or PBM measured at 10 sites in Canada during the years 2002–2011 were typically $< 5 \text{ pg/m}^3$ (Cole et al., 2014), corresponding with sample loadings of $3\text{--}9 \text{ pg}$. The corresponding median signal processing bias would be within -19% and -8% , respectively, based on the fit in Fig. 3a and within -12% and -8% , respectively, based on the fit in Fig. S5a.~~

Formatted: Superscript

Formatted: Superscript

Formatted: Superscript

~~My results demonstrate that s~~Signal processing uncertainties ~~can may~~ represent a significant fraction of the overall uncertainty in Tekran[®]-based atmospheric Hg measurements. I estimate that signal processing uncertainties in Hg loadings derived from the Tekran[®] method are within $\pm[14\% + 10.294 \text{ pg}]$ for the 2537A dataset and within $\pm[6\% + 0.21 \text{ pg}]$ for the 2537B dataset.

By comparison, non-signal processing related uncertainties in Tekran[®]-based GEM, TGM, and THg measurements were previously estimated to be on the order of ~10% to 15% for loadings typically <10 pg (Ambrose et al., 2011; Ambrose et al., 2013, 2015). Quantification of signal processing uncertainties in future applications of the Tekran[®] 2537 instrument will substantially improve the quality of the resulting measurements.

My results suggest that the performance of the Tekran[®] method may be improved with hardware modifications that increase the instrument's signal-to-noise ratio. The Tekran[®] method performed slightly better at low Hg loading when applied to the 2537B dataset than when applied to the 2537A dataset, ~~particularly at low Hg loading~~. The primary difference between the 2537A and 2537B instruments I tested is that the 2537B instrument was modified to improve its signal-to-noise ratio by replacing the ~~standard~~ sample cuvette and detector bandpass filter with a mirrored cuvette and an improved filter, respectively (Ambrose et al., 2013). ~~The cuvette and bandpass filter were obtained from Tekran[®] Instruments (Part Numbers 40-25105-00 and 40-25200-02, respectively).~~

It is possible that the performance of the Tekran[®] method can also be improved through modification of the method's integration parameters, though I tested only the default parameters (Table S1). Additionally, it is possible that measurement bias introduced by the Tekran[®] method can be made smaller by calibrating at loadings more similar to the loadings in the samples of interest. My results demonstrate a minor reduction in bias when measurements made with the 2537B instrument are calibrated at loadings that are >5 times higher rather than at loadings >10 times higher.

Measurement bias and precision can in principle both be improved by employing longer sample pre-concentration times and/or higher sample flow rates to achieve higher sample loadings.

Estimated signal processing uncertainties in Hg loadings derived from my methods range from within $\pm 0.053 \text{ pg}$ when the Hg thermal desorption peaks are defined manually to within $\pm[0.2\% + 0.053 \text{ pg}]$ (2537A dataset) and $\pm[2\% + 0.080 \text{ pg}]$ (2537B dataset) when Hg TD peak definition is fully automated. Biases in Hg loadings derived from my methods are lower by >80% than biases derived from the Tekran[®] method. Limits of detection for Hg decrease by 31 ~~to~~ ^{to} 88% when my methods are used in place of the Tekran[®] method.

Acknowledgements. Financial support for the Hg science component of NOMADSS was provided by the U.S. National Science Foundation, Award No. 1217010. Thanks to the Jaffe Research Group for helpful discussions during manuscript preparation. Thanks also to Franz Slemr for his constructive review.

Formatted: Superscript

Formatted: Font: Italic

References

- Ambrose, J. L., Gratz, L. E., Jaffe, D. A., Campos, T., Flocke, F. M., Knapp, D. J., Stechman, D. M., Stell, M., Weinheimer, A. J., Cantrell, C. A., and Mauldin, R. L. III: Mercury emission ratios from coal-fired power plants in the southeastern United States during NOMADSS, *Environ. Sci. Technol.*, 49, 10,389–10,397, doi:10.1021/acs.est.5b01755, 2015.
- 5 Ambrose, J. L., Lyman, S. N., Huang, J., Gustin, M. S., and Jaffe, D. A.: Fast time resolution oxidized mercury measurements during the Reno Atmospheric Mercury Intercomparison Experiment (RAMIX), *Environ. Sci. Technol.*, 47, 7285–7294, doi: 10.1021/es303916v, 2013.
- Ambrose, J. L., Reidmiller, D. R., and Jaffe, D. A.: Causes of high O₃ in the lower free troposphere over the Pacific Northwest as observed at the Mt. Bachelor Observatory, *Atmos. Environ.*, 45, 5302–5315, doi:10.1016/j.atmosenv.2011.06.056, 2011.
- 10 Cole, A. S., Steffen, A., Eckley, C. S., Narayan, J., Pilote, M., Tordon, R., Graydon, J. A., St. Louis, V. L., Xu, X., and Branfireun, B. A.: A Survey of Mercury in Air and Precipitation across Canada: Patterns and Trends, *Atmosphere*, 5, 635-668, doi:10.3390/atmos5030635, 2014.
- Gay, D. A., Schmeltz, D., Prestbo, E., Olson, M., Sharac, T., and Tordon, R.: The Atmospheric Mercury Network: measurement and initial examination of an ongoing atmospheric mercury record across North America, *Atmos. Chem. Phys.*, 13, 11339–11349, doi:10.5194/acp-13-11339-2013, 2013.
- 15 Gratz, L. E., Ambrose, J. L., Jaffe, D. A., Shah, V., Jaeglé, L., Stutz, J., Festa, J., Spolaor, M., Tsai, C., Selin, N. E., Song, S., Zhou, X., Weinheimer, A. J., Knapp, D. J., Montzka, D. D., Flocke, F. M., Campos, T. L., Apel, E., Hornbrook, R., Blake, N. J., Hall, S., Tyndall, G. S., Reeves, M., Stechman, D., and Stell, M.: Oxidation of mercury by bromine in the subtropical Pacific free troposphere, *Geophys. Res. Lett.*, 42, 10,494–10,502, doi:10.1002/2015GL066645, 2015.
- 20 Gustin, M. S. and Jaffe, D.: Reducing the uncertainty in measurement and understanding of mercury in the atmosphere, *Environ. Sci. Technol.*, 44, 2222–2227, 2010.
- Jaffe, D. J., Prestbo, E., Swartzendruber, P., Weiss-Penzias, P., Kato, S., Takami, A., Hatakeyama, S., and Kajii, Y.: Export of atmospheric mercury from Asia, *Atmos. Environ.*, 39, 3029–3038, 2005.
- Landis, M. S., Stevens, R. K., Schaedlich, F., and Prestbo, E. M.: Development and characterization of an annular denuder methodology for the measurement of divalent inorganic reactive gaseous mercury in ambient air, *Environ. Sci. Technol.*, 36, 3000–3009, 2002.
- 25 Lyman, S. N. and Jaffe, D. A.: Formation and fate of oxidized mercury in the upper troposphere and lower stratosphere, *Nature Geosci.*, 5, 114–117, doi:10.1038/ngeo1353, 2012.
- Obrist, D., Tas, E., Peleg, M., Matveev, V., Faïn, X., Asaf, D., and Luria, M.: Bromine-induced oxidation of mercury in the mid-latitude atmosphere, *Nature Geosci.*, 4, 22–26, doi:10.1038/NGEO1018, 2011.
- 30 Pandey, S. K., Kim, K.-H., and Brown, R. J. C.: Measurement techniques for mercury species in ambient air, *TrAC-Trend. Anal. Chem.*, 30, 899–917, doi:10.1016/j.trac.2011.01.017, 2011.

- Schroeder, W. H., Anlauf, K. G., Barrie, L. A., Lu, J. Y., Steffen, A., Schneeberger, D. R., and Berg, T.: Arctic springtime depletion of mercury, *Nature*, 394, 331–332, 1998.
- Schroeder, W. H., Keeler, G., Kock, H., Roussel, P., Schneeberger, D., and Schaedlich, F.: International field intercomparison of atmospheric mercury measurement methods, *Water Air Soil Poll.*, 80, 611–620, 1995.
- 5 Slemr, F., Weigelt, A., Ebinghaus, R., Kock, H. H., Bödewandt, J., Brenninkmeijer, C. A. M., Rauthe-Schöch, A., Weber, W., Hermann, M., Becker, J., Zahn, A., and Martinsson, B.: Atmospheric mercury measurements onboard the CARIBIC passenger aircraft, *Atmos. Meas. Tech.*, 9, 2291–2302 doi:10.5194/amt-9-2291-2016, 2016.
- Stratton, W. J. and Lindberg, S. E.: Use of a refluxing mist chamber for measurement of gas-phase mercury(II) species in the atmosphere, *Water Air Soil Poll.*, 80, 1269–1278, 1995.
- 10 Swartzendruber, P., Jaffe, D. A., and Finley, B.: Improved fluorescence peak integration in the Tekran 2537 for applications with suboptimal sample loadings, *Atmos. Environ.*, 43 (22–23), 3648–3651, 2009.
- Timonen, H., Ambrose, J. L., and Jaffe, D. A.: Oxidation of elemental Hg in anthropogenic and marine air masses, *Atmos. Chem. Phys.*, 13, 2827–2836 doi:10.5194/acp-13-2827-2013, 2012.
- Talbot, R., Mao, H., Scheuer, E., Dibb, J., and Avery, M. Total depletion of Hg⁰ in the upper troposphere-lower stratosphere, *Geophys. Res. Lett.*, 34 (L23804) doi:10.1029/2007GL031366, 2007.
- 15 Tekran Corporation: Model 2537A Ambient Mercury Vapour Analyzer User Manual, Rev. 3.01, Tekran Instruments Corporation, Toronto, Canada, 2006.
- Tekran Corporation: Model 2537B Ambient Mercury Vapor Analyzer User Manual, Rev. 3.10, Tekran Instruments Corporation, Toronto, Canada, 2007.
- 20 USEPA: Method 1631, Revision E: Mercury in water by oxidation, purge and trap, and cold vapor atomic fluorescence spectrometry, U. S. Environmental Protection Agency, Office of Water, Washington, DC, USA, 2002.

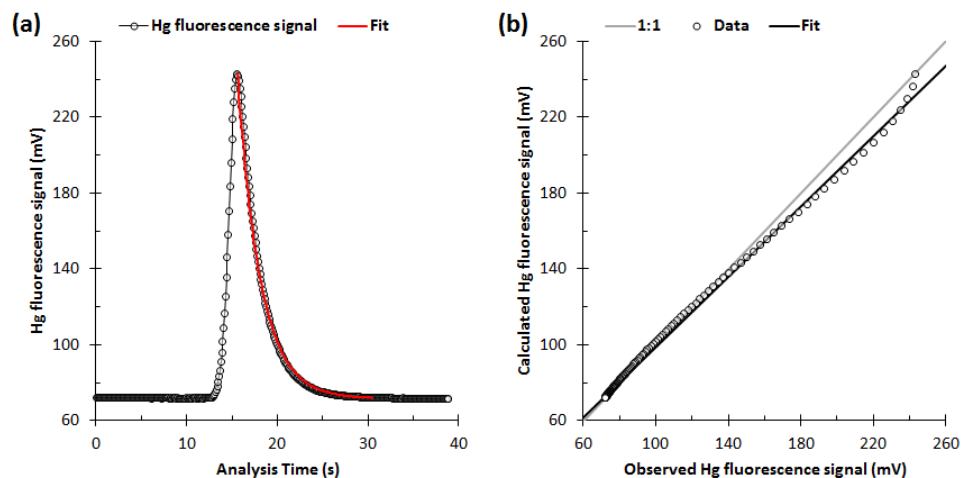


Figure 1: (a) Example Hg thermal desorption profile during a calibration gas analysis cycle on a Tekran® 2537A instrument. Also shown is the corresponding 150 point exponential Bisquare (unweighted) regression fit (eq. 1; $r^2 = 0.998$) used to derive the decay constant ($b = -0.041 \pm 0.004 \text{ ds}^{-1}$) during initialization of the VI's signal processing method. (b) Comparison between the calculated (fit) and observed Hg atomic fluorescence signal values in panel (a). The slope and intercept of the linear fit are 0.921 ± 0.007 and $7.2 \pm 0.8 \text{ mV}$, respectively (uncertainties are 95% confidence intervals). (Analogous results obtained with a 2537B analyzer are shown in Fig. S4 in the Supporting Information.)

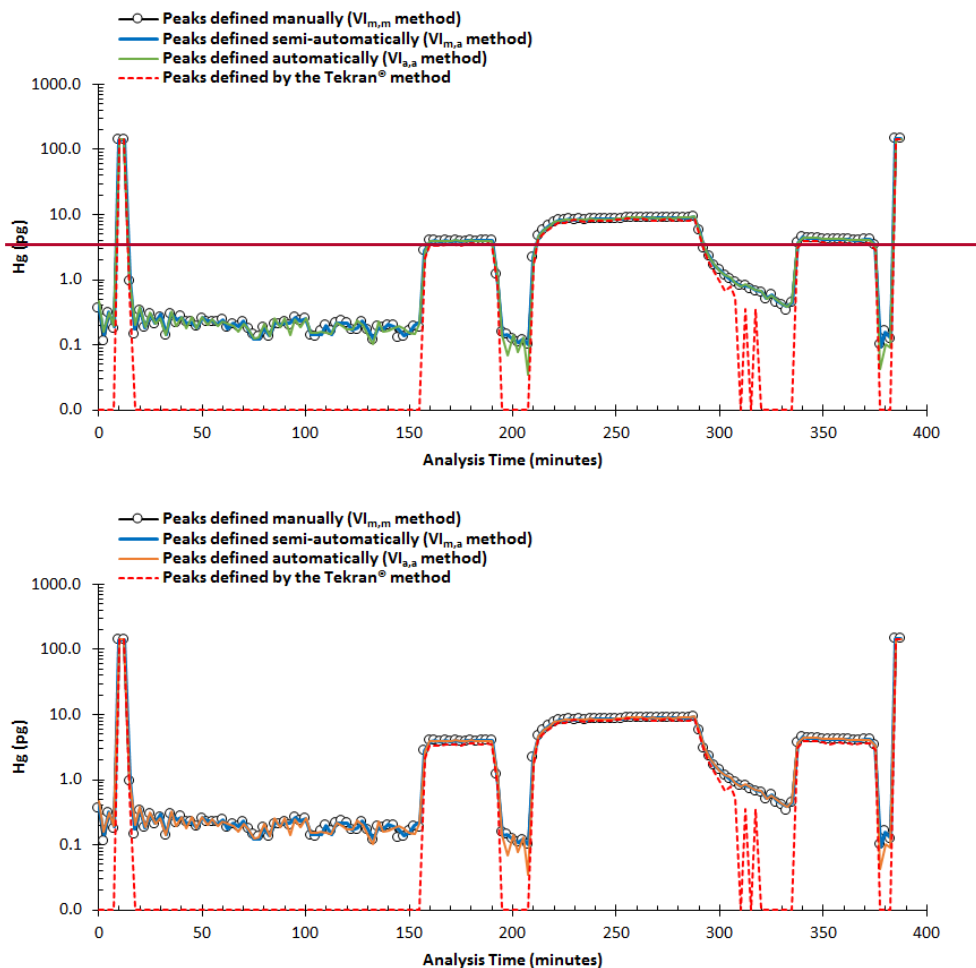


Figure 2: Test dataset collected with a Tekran® 2537A instrument, represented as Hg loadings derived from my VI-based manual, semi-automated, and automated Hg thermal desorption peak height determination methods (the VI_{m,m}, VI_{m,a}, and VI_{a,a} methods, respectively), and by the Tekran® method. Peaks not detected by the Tekran® method are assigned a value of 0.01 pg. The two pairs of data points at >100 pg correspond with calibration gas analysis cycles (SPAN samples). I use the first pair of SPAN samples to initialize the VI (as described in Sect. 2.1.2). I use response factors calculated from the second pair of SPAN samples (and the

preceding pair of blanks) to calculate Hg loadings for all other samples in the dataset. The mean value of the baseline standard deviation, σ_M (defined in Sect. 2.1.1), is ~ 0.03 mV (equal to ~ 0.03 pg). The corresponding estimated lower-limit f value (mean $\pm 2\sigma$; Sect. 2.1.3) is $1.86(1) \times 10^{-4}$.

5

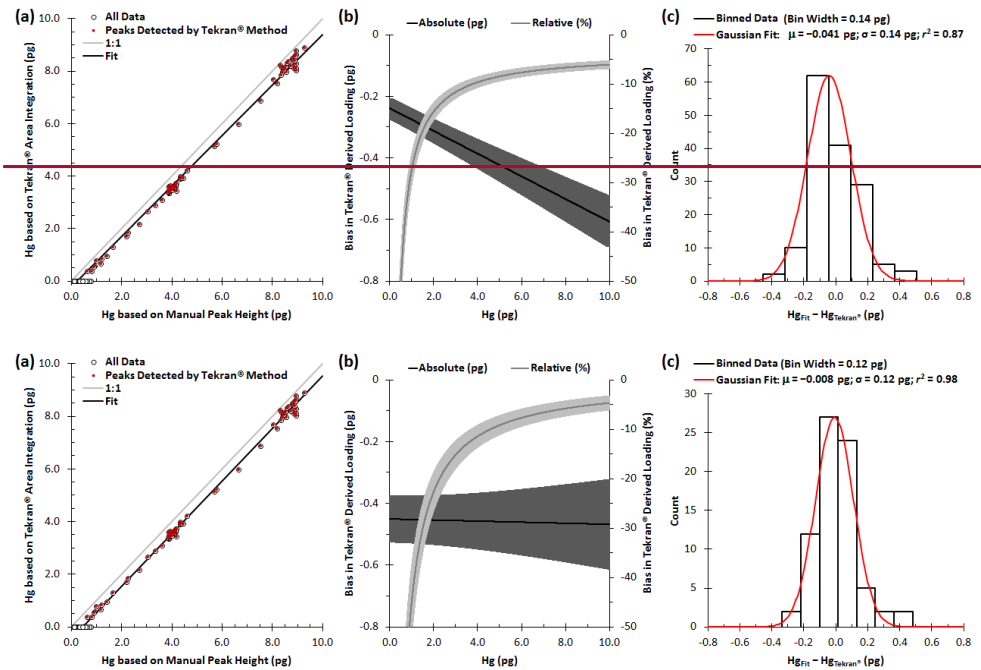
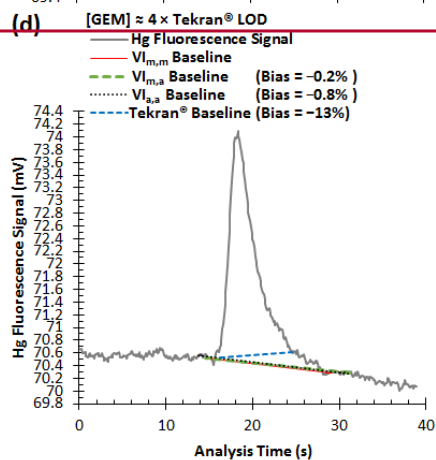
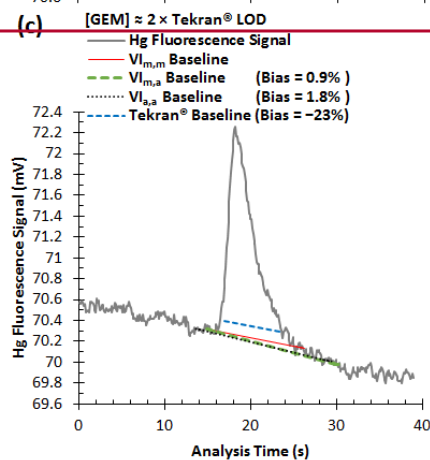
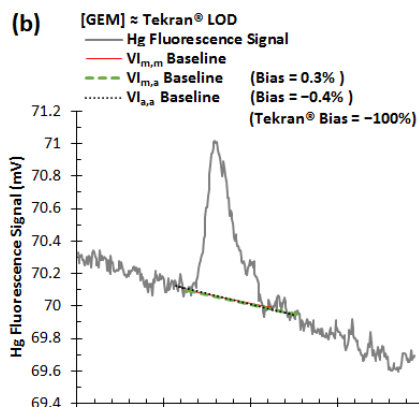
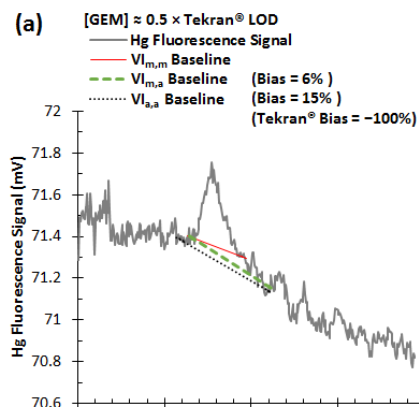


Figure 3: (a) Comparison of Hg loadings derived from measurements made with a Tekran® 2537A instrument using the Tekran® method and my VI-based Hg thermal desorption peak height determination method (dataset shown in Fig. 2, excluding SPANs), with the peaks defined manually (the $VI_{m,m}$ method). The equation of the linear regression is $y = 1.000963(17)x - 0.4524(83)$ pg ($r^2 = 0.9978$, $n = 74152$). The fit excludes data derived from peaks not detected by the Tekran® method (represented by the open symbols). (b) Absolute and relative biases in the Tekran®-derived Hg loadings, based on the fit in panel (a). The grey bands represent propagated uncertainties (95% confidence intervals)

in the fit parameters. (c) Distribution of residuals from panel (a), including only data derived from detected peaks. (Analogous results for the 2537B instrument are presented in Fig. S5 in the Supporting Information.)



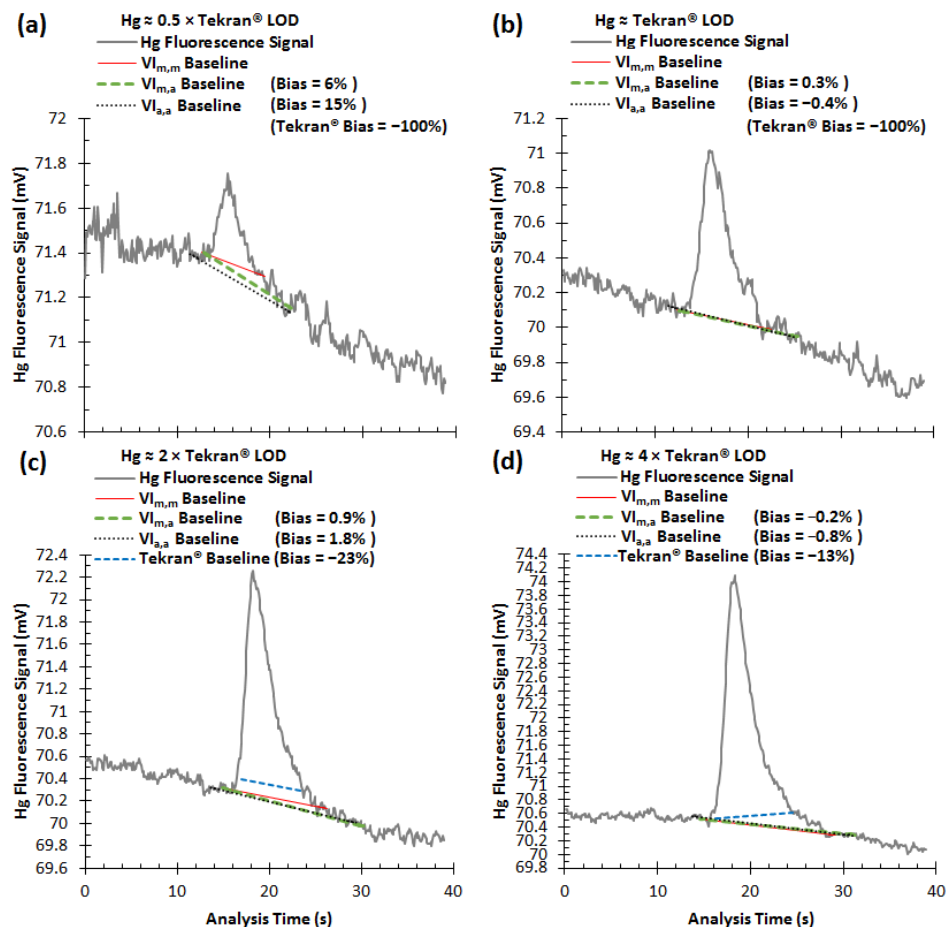


Figure 4: Comparison of Hg fluorescence baselines calculated by applying manual ($VI_{m,m}$ method), automated (Tekran® and $VI_{a,a}$ methods), and semi-automated ($VI_{m,a}$ method) Hg thermal desorption peak definition methods to samples with Hg loadings of approximately (a) 0.5, (b) 1, (c) 2, and (d) 4 times the limit of detection of the Tekran® method (~0.8 pg). Measurements were made with a 2537A instrument. Biases in the Hg loadings derived from the Tekran® method and the VI-based methods are indicated. Biases are expressed relative to the loadings derived from the $VI_{m,m}$ method and are negative when the $VI_{m,m}$ -based loadings are higher. The Tekran® baselines are missing from panels (a) and (b) because the peaks ~~are~~ were not detected by the Tekran® method.

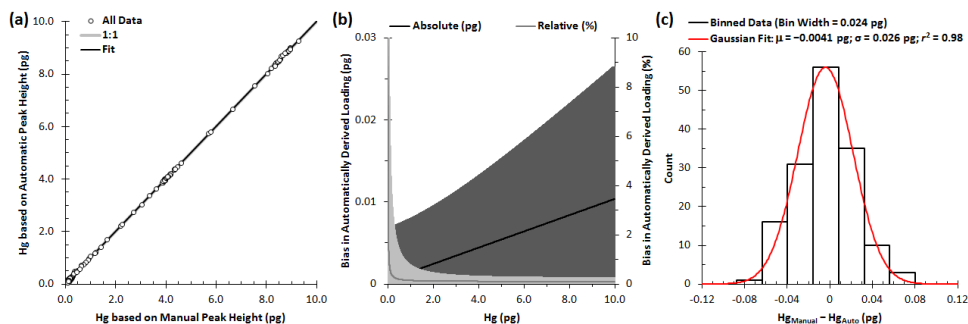


Figure 5: (a) Comparison of Hg loadings derived from measurements made with a Tekran® 2537A instrument using my VI-based automated and manual peak height determination methods (the VI_{a,a} and VI_{m,m} methods, respectively; dataset shown in Fig. 2, excluding SPANs). The equation of the linear regression is $y = 1.001(1)x + 0.000(6)$ pg ($r^2 = 0.99992$, $n = 152$). (b) Absolute and relative biases in the VI-based Hg loadings, based on the fit in panel (a); grey bands represent propagated uncertainties (95% confidence intervals) in the parameters of the fit in panel (a). (c) Distribution of residuals from panel (a).

Table 1. Bias in Hg loadings derived by applying automated and semi-automated Hg atomic fluorescence signal processing methods to measurements made with a Tekran® 2537A instrument.

Method	Hg (pg, ng/m ³) ^{a,b}						
	7.5, 1.5	3.75, 0.75	2.5, 0.5	1.25, 0.25	0.5, 0.1	0.25, 0.05	0.125, 0.025
Tekran ^{®c}				-36.23 ± 6.3	-100 ^f	-100 ^f	-100 ^f
VI _{a,a} ^d	-67.4 ± 20.9	-121 ± 21	-184 ± 31	0.2 ± 0.5	0.3 ± 1.2	0.6 ± 2.3	1 ± 5
VI _{m,a} ^e	0.15 ± 0.08	0.2 ± 0.1	0.3 ± 0.1	0.5 ± 0.3	1.0 ± 0.6	2 ± 1	4 ± 2

^aBias values for the Tekran® and VI_{a,a} methods are calculated from the equations of the linear regressions in Figs. 3a and 5a, respectively. Bias values for the VI_{m,a} method are similarly calculated from the linear regression equation given in Table S4 in the Supporting Information (“Standard” configuration). All bias values are expressed relative to Hg loadings derived by processing the data using manual peak definition (the VI_{m,m} method).

^bHg loadings are also expressed in terms of concentrations under the typical Tekran® operating parameters.

^cTekran® operating and peak integration parameters are defined in Table S1.

^dMy VI-based peak height determination method, with peak start and end times determined automatically (VI_{a,a}).

^eMy VI-based peak height determination method, with peak start times determined manually and peak end times determined automatically (VI_{m,a}).

^fFor Hg < the estimated 0.8 pg Tekran® limit of detection, the true bias is -100%. For clarity, the true bias is substituted for the calculated values.

Table 2. Estimated signal processing uncertainties and Hg limits of detection (LODs) achieved with the Tekran® peak integration method and my VI-based peak height determination methods as applied to measurements made with Tekran® 2537A and 2537B instruments.

Method ^a	Signal Processing Uncertainty ^b	Hg LOD (pg)
2537A Dataset ^c		
Tekran®	±[14% + 10.294 pg]	0.80 ^d
VI _{m,m}	±0.053 pg	0.10 ^e
VI _{a,a}	±[0.2% + 0.053 pg]	0.12 ^e
VI _{m,a}	±[0.2% + 0.053 pg]	0.10 ^e
2537B Dataset ^f		
Tekran®	±[6% + 0.21 pg]	0.20 ^d
VI _{m,m}	±0.030 pg	0.085 ^g
VI _{a,a}	±[2% + 0.080 pg]	0.13 ^g
VI _{m,a}	±[0.6% + 0.030 pg]	0.10 ^g

^aThe "Tekran®" method is the Tekran® analyzer's internal automated Hg thermal desorption peak integration method, parameterized as indicated in Table S1 in the Supporting Information. The "VI_{m,m}", "VI_{a,a}", and "VI_{m,a}" Hg TD peak height determination methods were developed in this work and are described in Sect. 2.1.2. The operating parameters of the Tekran® analyzers are presented in Table S1.

^bEstimated as described in Sect. 3.1.4.

^cThe 2537A dataset is shown in Fig. 2.

^dEstimated as the highest Hg loading derived from the VI_{m,m} method for samples for which the Tekran® method failed to detect the Hg TD peak.

^eEstimated as twice the standard deviation of blank loadings (*n* = 62).

^fThe 2537B dataset is shown in Fig. S3 in the Supporting Information.

^gEstimated as twice the standard deviation of blank loadings (*n* = 37).

Formatted: Superscript

Formatted: Superscript